Field experimental evidence on the incentive effects of agricultural contracts on real-effort output in Ethiopia*

Niels Kemper†

March 2012

Abstract

This study presents results of a real-effort field experiment on the real-effort output induced by agricultural contracts common in rural Ethiopia. The design of the experiment resembles sorting processes archetypical for agrarian societies which are relatively abundant in labor and relatively scarce in capital. Drawing on a random sample of $n = 500$ households, I find that fixed-rent and owner contracts induce the highest real-effort output as compared to the wage contract (in line with basic economic theory). Surprisingly, real-effort output from sharecropping and wage contracts cannot be statistically distinguished. I conjecture that this is due to a real life spillover into the experiment impairing the incentive provision by the sharecropping contract.

Keywords: Agricultural contracts; Ethiopia; Real-effort experiment.

JEL classification codes: (JEL C9, J3, M5)

1 Introduction

There are four polar cases in the decision of how to contract land and labor and hence the claim to yields from land: wage labor, sharecropping, fixed-rent

*I would like to express my gratitude to the Stiftungsfond of the Deutsche Bank for funding this field research and to Rainer Klump from the Goethe-University in Frankfurt, Germany, for writing the research proposal with me. I am also grateful to Tekie Alemu and Tadele Ferede from the School of Economics at the Addis Ababa University, Ethiopia, for their intellectual, logistic and administrative support of the research project. I am also indebted to my field supervisor Getachew Andane as well as my enumerators Kassahun Alemayehu, Alem Brhanemeskel, Sisay Addis, Yeshi Mulatu, Fetene Belayhum, Ermias Samuel, Hibiret Tsehay, Daniel Mekonnen and Awoke Yalew for their dilligent and thorough field work. This research also benefited from seminal discussions on real-effort experiments and contract theory with Myriam Hadnes, Matthias Heinz and Heiner Schumacher from the Goethe-University in Frankfurt and Markus Frölich and Andreas Landmann from the University of Mannheim, Germany.

†University of Mannheim, Department of Economics, L7, 3-5, DE-68131 Mannheim, Germany, +49-(0)621-181-1805, niels.kemper@uni-mannheim.de.

and owner cultivation, of which sharecropping is predominantly reflected on in the literature. The term sharecropping comprises a contractual relationship between a landlord and a tenant allowing the latter to use the land of the former in return for a share of the produced crops on this land. The discussion on the (in)efficiency of sharecropping is almost as old as the discipline of Economics. Adam Smith and John Stuart Mill were early proponents of the Englisch fixed-rent leasehold tenancy vis-a-vis the French *metayage* (share tenancy), as in the former contractual arrangement tenants have stronger incentives to invest more effort. Alfred Marshall argues along these lines in *Principles of Economics*, formally showing that sharecropping results in too low labor supply as the tenant equates the share of output times the marginal productivity of labor to the disutility of work, while Pareto optimality requires the marginal productivity of work to equal the marginal productivity of labor.

Empirically, evidence on the so-called Marshallian inefficiency is mixed. In Ethiopia, Pender and Fafchamps (2006) find no differences in input and output intensities for owned and sharecropped plots, but Teklu and Lemi (2004), Holden and Ghebru (2006), Ghebru and Holden (2006), Deininger et al. (2008) do find (substantial) differences and Deininger et al. (2011) find differences unless for sharecropping contracts which involve cost sharing (which are, however, rather uncommon). Measuring the pure incentive effect of agricultural contracts is complicated in a number of ways. The choice of a sharecropping over a fixed-rent contract, for instance, may be determined by a risk-sharing motive (Cheung, 1969), a trade-off between risk-sharing and incentive provision (Stiglitz, 1974) and screening with self-selection into contracts (Hallagan, 1978, and Newbery and Stiglitz, 1979). This confounds the empirical analysis of contract choices as determinants like individual risk attitudes, exercised effort and ability, for instance, are difficult to measure an thus to control for.

This study takes a framed field experimental approach to study the incentive effects of agricultural contracts typical for developing countries. The experimental setting allows to vary contract treatments independently of the confounding factors above and thus to elicit the pure incentive effect of contracts. In particular, the framed field experimental approach allows (i) to measure the incentive effects of contracts with a task natural to subjects, (ii) in their naturally occurring environment, (iii) when behavior is not free of context.

In different settings, the incentive provision of contracts has been studied in a number of natural field experiments. Lazear (2000) analyzed a shift from wage rate to piece rate compensation at the Safelite Glass cooperation, a large US auto glass company. Hamilton et al. (2003) analyse the shift from individual to group piece rate compensation in a garment manufacturing facility in the US. Shearer (2004) estimates the gain in productivity realized when workers are paid piece rates rather than fixed wages within a tree-planting firm. Bandiera et al. (2005) compare a shift from relative incentive schemes (worker's pay depends on the ratio of individual to average productivity) to piece rate compensation in a leading UK based fruit farm. While the approach taken in these studies is most suitable to analyze worker's response to compensation schemes in industrialized production process, they provide limited scope for learning in developing

countries where the majority of the workforce is found in the agricultural sector with highly decentralized, small scale production.

In a lab experiment, Dohmen and Falk (2011) hold constant all relevant factors which may determine contract choice. They compare productivity outcomes under piece rate, tournament or revenue-sharing compensation scheme after the self-selection of workers in the respective contracts. In another lab experiment Abeler et al. (2011) exogenously vary piece rate and fixed wage compensation schemes to study the effect of expectations on effort provision. While these studies are not impaired by sorting (that is, self-selection of workers into contracts) which may affect natural field experiments, they are of no avail in situations in which real-word actors behave different from students in the lab. In particular, they are of no avail in situations where socio-economic context matters for the behavior of participants.

The analysis in this paper is based on data collected at different study sites in rural Ethiopia. I designed a real-effort task resembling sorting processes that occur in agricultural production. Given that the economy in rural Ethiopia is relatively abundant in labor and relatively scarce in capital, these sorting processes are quite common (see section 3.2 for examples and the design of the real-effort task). The real-effort task involves the sorting of haricot beans of different colors in a given time period. The experiment consists of 25 sessions with 20 participants each. Agricultural contracts were randomly assigned to sessions (the wage treatment to 7 sessions, the sharecropping, fixed-rent treatment and owner treatment to respectively 6 sessions). The payoff of the participants depended on the terms specified in the respective contract and the participant's real-effort output during the experiment (except for the wage treatment were the payoff was held fixed for all levels of real-effort output). As opposed to many other studies on the

In the empirical analysis the wage treatment serves as the baseline case for the estimation of treatment effects (see section 2.2 for the definition of the treatment effects). Identification issues arising from the fairly low number of randomization per treatment are discussed and results are shown to be robust to a number of model specifications. The main results (see table 8, appendix A) are as follows:

1. The fixed-rent treatment as compared to the wage treatment increases real-effort output between 8.9 and 9.6 percent.

2. The owner treatment as compared to the wage treatment increases real-effort output between 6.4 and 7.2 percent.

3. Given that I chose the wage payment in the wage treatment to be sufficiently low, it is surprising to find that real-effort in the sharecropping treatment cannot be statistically distinguish from the wage treatment (see section 3.2 on details on the implementation of the payoff schemes).

Semi-structured interviews imply discontent with real life sharecropping contracts among participants across almost all study sites (see section 5). The main

reason are the terms of the sharecropping contracts. I show that the subgroup of real life sharecroppers behaves different from non-sharecroppers for both the sharecropping and the fixed-rent treatment, which implies a real life spillover into the experiment.

Looking at correlations between real contracts chosen during the last cropping season and per-capita expenditure levels of our sample households (see table 1 in appendix A) implies that there might be a welfare dimension to agricultural contract choice in real life. While sharecropping households have a monthly per-capita expenditure level being 78 Birr lower than for other households, households with fixed-rent contracts have a per-capita expenditure level that is 127 Birr higher as compared to other households. This is substantial given an average monthly per-capita expenditure level of 727 Birr. Given that rural Ethiopian households typically make a living on subsistence agriculture, this implies a relationship between agricultural contract choices and poverty.

This paper is structured as follows: Section 2 provides the conceptual framework, the testable hypotheses and specifies the treatment effects estimated in the experiment. Section 3 describes the experimental design and procedure. Section 4 discusses issues related to identification of and inference on the estimated coefficients and presents the result from the experiment. Section 5 concludes. The appendix contains tables, variable definitions and pictures from the set-up of the experiment.

2 Conceptual framework

2.1 Agricultural contracts and effort

This section introduces contracts typical for Ethiopian agriculture and derives some testable hypotheses based on very simple economic reasoning. Let the following payoff functions describe a class of contracts which polar cases are typical for landlord and tenant relationships in developing countries (similar to Stiglitz, 1974):

$$y_T = \alpha * grams + \beta \tag{1}$$

$$y_L = (1 - \alpha) * grams - \beta \tag{2}$$

$$grams = e^{\gamma} \tag{3}$$

where y_T and y_L denote the payoff of tenant and landlord, $0 \leq \alpha \leq 1$ the tenant's share, $\beta > 0$ a wage and $\beta < 0$ a fixed-rent payment. The variable e denotes extra real-effort (which comes in addition to some unspecified but individual-specific normal effort) and $grams$ denotes extra real-effort output. The elasticity of $grams$ with respect to e is captured by $0 < \gamma < 1$.[1] This set

[1] This implies decreasing returns to effort. The real-effort task in the experiment has decreasing returns by construction.

of equations contains the following contracts as polar cases:

- Wage labor for $\alpha = 0$ and $\beta > 0$
- Sharecropping for $\beta = 0$ and $0 < \alpha < 1$
- Fixed-rent for $\alpha = 1$ and $\beta < 0$

We focus on polar cases as mixed contracts (e.g. $0 < \alpha < 1$ and $\beta \neq 0$) are rarely observed empirically. Otsuka et al. (1992) criticize that early sharecropping models have treated land and labor contracts as separate (e.g. Stiglitz, 1974, looks only at sharecropping and fixed-rent contracts) and have been ignorant of owner cultivation. For the sake of completeness, this makes the case for a fourth special case, a one-party contract:

- Owner cultivation for $\alpha = 1$ and $\beta = 0$.

Assuming that tenants equates marginal utility of extra output to marginal cost of extra effort $c(e)$ to (1), than the tenant will supply effort as follows:

$$\gamma \alpha e^{\gamma - 1} = c'(e) \tag{4}$$

In marginal terms, only the contract parameter α matters for the supply of effort. Given that for the wage contract $\alpha = 0$, for the sharecropping contract $0 < \alpha < 1$ and for fixed-rent and owner contract $\alpha = 1$, it follows that fixed-rent and owner contract induce the same effort (as a fixed-rent is just a parallel downward shift of the production function). The effort induced by the sharecropping contract is lower, because multiplying the production function by a constant $0 < \alpha < 1$ changes the slope of the production function. And given that $\alpha = 0$, the wage contract induces no extra effort at all.

2.2 Experimental approach to comparing effort

Given that real-effort is unobserved, I estimate the effect of randomly assigned contract treatments on real-effort output from the following population regression:

$$grams_{ij} = \delta_0 + \delta_1 c_{j1} + \delta_2 c_{j2} + \delta_3 c_{j3} + \delta_4 c_{j4} + w_{ij} \tag{5}$$

where c_{j1}, c_{j2}, c_{j3} and c_{j4} are binary variables equal to one if a particular contract is assigned to session $j = 1, 2...25$ and zero otherwise and $grams_{ij}$ denotes the real-effort output by individual $i = 1, 2...20$ in session j and w_{ij} is a random disturbance term for individual i in session j. For the causal evaluation of the incentive structure of the different contracts I have to find a proper counterfactual to compare potential outcomes. Two approaches are obvious.

First, in an experimental setting it is indeed possible to (subsequently) assign different contracts to the same individual and see how effort varies. While this has the merit of eliminating unobserved individual fixed effects by comparing

the outcomes for the same individual under different treatments, it introduces learning into the experiment difficult to account for. This is empirically hard to disentangle from the pure incentive effect of the contract (even when the order of assignment is randomized for individuals). Therefore I take a different approach.

Second, I compare real-effort output for contract treatments randomly assigned to different subjects in which a subgroup of participants serves as counterfactual. Given that working under no contract is an infeasible counterfactual for working in a contract (because of the binding individual participation constraint participants do not exercise effort if they are not incentivized), I specify one of the contracts as reference contract. While this could be any of the contracts, I take the wage contract as reference contract.

Thinking about this in terms of potential outcomes and assuming that the wage treatment is independent of the counterfactual outcome for any of the other contract treatments, for instance, c_2, then we can write

$$E(grams_{0ij} \mid c_{j1} = 1, c_{j2} = 0, c_{j3} = 0, c_{j4} = 0) = \delta_0 + \delta_1 \tag{6}$$

assuming that δ_1 is additive and constant. And given that the effect of c_2 is additive and constant

$$E(grams_{1ij} \mid c_{j1} = 1, c_{j2} = 1, c_{j3} = 0, c_{j4} = 0) = \delta_0 + \delta_1 + \delta_2. \tag{7}$$

Then δ_2 can be interpreted as average treatment effect for the sharecropping treatment:

$$\begin{aligned} \delta_2 &= E(grams_{1ij} \mid c_{j1} = 1, c_{j2} = 1, c_{j3} = 0, c_{j4} = 0) - \\ E(grams_{0ij} &\mid c_{j1} = 1, c_{j2} = 0, c_{j3} = 0, c_{j4} = 0) = ATE_{share}. \end{aligned} \tag{8}$$

Similarly, the interpretation for δ_3, the fixed-rent treatment, is

$$\begin{aligned} \delta_3 &= E(grams_{1ij} \mid c_{j1} = 1, c_{j2} = 0, c_{j3} = 1, c_{j4} = 0) - \\ E(grams_{0ij} &\mid c_{j1} = 1, c_{j2} = 0, c_{j3} = 0, c_{j4} = 0) = ATE_{rent} \end{aligned} \tag{9}$$

and for δ_4, the owner treatment, is

$$\begin{aligned} \delta_4 &= E(grams_{1ij} \mid c_{j1} = 1, c_{j2} = 0, c_{j3} = 0, c_{j4} = 1) - \\ E(grams_{0ij} &\mid c_{j1} = 1, c_{j2} = 0, c_{j3} = 0, c_{j4} = 0) = ATE_{own}. \end{aligned} \tag{10}$$

To implement the regression in (5) in the data sample at hand I set $c_{j1} = 0$ to estimate the ATEs for δ_2, δ_3 and δ_4.

3 Experimental design and procedure

3.1 Data and sampling

The study sites were selected at three stages. First, we defined the sample universe to consist of households in Amhara and Oromia, the two economically most active and most populated regions of Ethiopa, and respectively sampled two districts (Woredas) in each region (Gozamen and Bahir Dar Zuria in Amhara, Adaa and Girar Jarso in Oromia). Districts are composed of peasant associations (Kebeles), which are the smallest unit of administration in Ethiopia. We randomly selected Kebeles within Woredas (5 in Adaa, 5 in Girar Jarso, 8 in Gozamen and 7 in Bahir Dar Zuria).

Sampling by clusters in this three-stage design generates a sample in which households are not randomly distributed in terms of geography but geographically grouped. Given varying population sizes of the clusters, households are sampled with unequal selection probabilities. In the empirical analysis below I consider the sampling design by presenting population-weighted regressions next to unweighted regressions to check whether the empirical results can be considered representative for the underlying population.

There is census data on the different Ethiopian Kebeles which, however, as the authorities in Addis Ababa told us, is not always up to date.[2] Hence, for the sampling of households at the Kebele level we drew on registers kept by Kebele chiefs. These registers are widely held to be up to date as claims to land are documented in this register. Randomization at the Kebele level was done by dividing the N households living in the Kebele through 22 (for 20 participants and 2 replacements in case sampled participants do not show up) and then inviting every $N/22$ household head from the list. It was decided to invite the head rather than a randomly selected household member, because the experiment was accompanied by a detailed quantitative and qualitative household survey and household heads were considered the most appropriate person to get information from. The sampling at the Kebele level was done by an enumerator who travelled to the study sites one day before the data collection to facilitate the invitation of households.

The data collection took place in April and May 2011 before the start of the weeding season (which coincides with the start of the main rainy season called *Meher* in June). Therefore household heads could spare time to participate in our data collection (upon invitation they were told that they could earn an unspecified amount of money during a research visit from a team from the Addis Ababa University). The enumerators had to refer to the replacement household heads only in a very few cases.

The different sessions comprised experimental, observational and qualitative data collection. A typical session looked as follows. We arrived in the village and met local organizers. Then we met with sampled participants in a venue such as communal centers or health centers.

[2] The third and last Population and Housing Census so far was conducted in 2010.

The whole session consisted of two experiments (the real-effort game and a risk attitude experiment as in Tanaka et al. (2010)) and a 20 pages survey questionnaire combining quantitative data collection with semi-structured interviews. After enumerators introduced themselves, a text was read out to the participants. The formulations were the same across all sessions and were read out by the same enumerator and were taken (and slightly paraphrased) from the the scripts provided at the *Experiments in 15 Small-Scale Societies* website[3]. Participants received a 20 Birr show-up fee (roughly 1.20 USD at the time of data collection). They were told that they can quit the experiments anytime if they fell uncomfortable with it. However, all of the participants chose to accomplish the task.

Afterwards the first experiment was conducted, then the second. To avoid systematic spillover from one experiment into the other, we randomized the order of the experiments across all sessions (see section 4.5 for a test on this).

The enumerators carefully explained the real-effort experiment. They showed how to exercise the task and how not to exercise the task and participants answered control questions to demonstrate that they understood the task. They also told participants how they were paid. Explaining and conducting the real-effort experiment usually took more than 60 minutes. In all sessions the enumerators first conducted the experiments and then the questionnaire. Completing the data collection for a session typically took the better half of the day.

The data set, Stata do-files for the reproduction of the empirical results and the survey instruments can be downloaded from the internet. See Appendix C for further details.

3.2 The real-effort task and the random assignment of contract types

The sorting task used in the real-effort experiment resembles agricultural production processes archetypical for Ethiopia. While it is easy to come up with numerous examples for sorting in agriculture in general (it is particularly easy in an economy such as Ethiopia which is relatively abundant in labor and relatively scarce in capital), I focus on some examples from food production.

Indigenous cereal agriculture in Ethiopia has existed as long as 13000 B.C. according to some estimates. Climatic and geographic conditions in the northeast of the Ethiopian plateau created a suitable environment for wild grains people started to utilize and domesticate intensively at the end of the Pleistocene. Early domesticated grains were wheat, barely and teff. Cereal-based cultivation of these grains has persisted from ancient times up to the day.[4] (de Wet, 1977, Ehret, 1979).

[3] See http://www.hss.caltech.edu/~jensming/roots-of-sociality/ (Accessed on Feb 15th, 2011).

[4] In my sample they roughly amount to 50 percent of the land under cultivation: 20.3 percent, 12.4 percent, 6.0 percent and 11.6 percent for respectively white teff, black and mixed teff, barley and wheat.

These crops share common characteristics in processing. After harvesting, crops are threshed, winnowed and sieved. Sorting processes occur along the way. For instance, in teff processing raw straw is removed during threshing, during winnowing grain is separated from fine straw and chaff and during sieving grain is separated from guardies (D'Andrea et al., 1999). There are more examples of sorting processes typical for other Ethiopian regions, but not necessary the regions where I conducted the field research.

First, coffee is Ethiopia's main export crop. Dry processing of coffee beans is common. The first step in processing coffee is cleaning, which involves winnowing done by hand using a sieve. In the process harvest cherries are sorted and cleaned from dirt, soil, twigs and leaves. Unwanted cherries are picked out from the top of the sieve and removed.

Second, fruit plantations such as Avocado, banana, citrus, grape, pineapple, papaya, mango, peach, apple and vegetables are also quite common in Ethiopa. Pests of fruit crops is a major problem in all areas of fruit production as fruit flies are attracted by ripening or fermenting fruit. One way of dealing with fruit flies is the sorting and separating of ripe and unripe fruits after harvest.

I designed a real-effort task resembling sorting processes like these. During the experiment participants had to sort three different types of haricot beans (each having a different color). To be more precise: Participants had to pick one particular haricot bean from a piling up of the three types of beans. Performance was measured as the weight of beans participants managed to separate from the piling up. Depending on the session, the payoff for the real-effort task was done with respect to one out of four randomly assigned contract treatments. The payoff for performance for the different contracts were determined as follows. In the wage treatment participants were compensated with the fixed amount of 20 Birr for doing the real-effort task:

$$y = 20Birr \tag{11}$$

In the sharecropping treatment the compensation of participants is a fraction of real-effort output. In the experiment participants receive half of 4 Birr for additional 10 grams of real-effort output. An equal split of output between landlords and tenants is commonly observed in real sharecropping arrangements. The payoff for this treatment is:

$$y = (grams * 4Birr/10g) * 0.5 \tag{12}$$

where $grams$ is the real-effort output exercised in the real effort task. In the fixed-rent treatment the compensation of participants is reduced by the fixed-amount of 20 Birr. Put differently, participants may keep everything in excess of 20 Birr. The payoff is:

$$y = (grams * 4Birr/10g) - 20Birr \tag{13}$$

Finally, in the owner treatment participants are compensated with respect to their real-effort output without any deduction. The payoff for this treatment is:

$$y = (grams * 4Birr/10g) \tag{14}$$

These payment schemes were not chosen arbitrarily. During preparatory sessions with enumerators and students, the wage treatment yielded an average effort between 110 and $115grams$. I normalized the payoff for contract c_1, c_2 and c_3 clearly below that value to avoid that inappropriate scaling of the wage contract influences the exercised effort. That is, I normalize the payoff of c_1, c_2 and c_3 at a low level of real-effort output: For $grams = 100g$ the payoff for the respective contract treatments is in all three cases $20Birr$. The payoff of the owner treatment is $40Birr$.

Table 2 in appendix A shows some descriptive statistics for the contract treatments. There were 25 sessions in total. Each session had 20 participants. The sharecropping, fixed-rent and ownership contract treatment were each assigned 6 times, the wage contract 7 times. Two out of three haricot beans were randomized across contract treatments (yellow and white beans, but not red beans). Yellow and white beans were virtually alike in size, but white beans weighted a little more. Red beans were bigger and heavier than yellow and white beans. Therefore they were not used in the real-effort task. Accordingly the sharecropping, fixed-rent and ownership contract treatment were played 3 times each with white and a yellow beans each. The wage contract treatment was played 4 times with the white beans and 3 times with the yellow beans. On average, participants earned $20Birr$ in the wage treatment, $26.43Birr$ in the sharecropping treatment, $39.78Birr$ in the fixed-rent treatment and $58.43Birr$ in the owner treatment.

3.3 Experimental procedure

The basic means for conducting the real-effort experiment were two buckets, beans of three different colors and a weight scale (see picture 1, 2 and 3 in appendix E for a close-up of the beans, the equipment needed to conduct the experiment and the seating order during the experiment).

The setup of the experiment was as follows. Each participant received two buckets:

- A flat, wide bucket (blue) with a diameter of 22.1 centimeters. At the beginning of the experiment this bucket was filled with beans in three different colors (red, yellow, white). Each bucket contained 700 grams of beans of each color.

- A raised, narrow bucket (red) with a diameter of 16.3 centimeters. At the beginning of the experiment this bucket was empty.

All flat and all raised buckets respectively had the the same size, same shape and contained exactly the same weight of red, yellow and white beans. *For the real-effort task, I asked participants to collect as many beans of a particular color as they can and put them into the red bucket within a 15 minutes time*

period. Time was stopped with a stopwatch, the stopwatch was the same over all sessions. It was checked that all participants started at the same time and finished at the same time.

My biggest concern was to establish independence in the exercised effort of each participant to make sure I measure the pure incentive effect of each contract treatment rather than peer effects (or something else). I designed the experimental protocol to prevent participants from influencing each other. Doing so, the enumerators took the following steps: First, they arranged participants to sit in a circle with adequate distances between them. The experiment was conducted outside on meadows to allow for enough space. Second, to make sure that participants do not see each other performing the task the enumerators asked participants to turn around such that their backs were facing the inside of the circle (because participants might increase their effort if they watch other participants exercise high effort - or vice versa). Third, to make sure that participants do not hear each other performing the task the enumerators covered the bottom of the bucket with a paper towel such that beans did not produce a sound when they were dropped into the bucket (because participants might increase their effort if they hear other participants exercise high effort - or vice versa). Further, the enumerators turned on music on portable speakers. However, given that the paper towel worked well this was more useful in entertaining participants during the task rather than preventing them from hearing each other. The enumerators played the same song (twice) during each session.

After the seating was in order they arranged the buckets for the participants. To make sure that they do not miss out on left-handers the enumerators asked participants which hand they use to pick-up things (besides food as Muslims never use the left hand for eating). For right-handers the empty bucket was put to their left and for left-handers to the right. The flat bucket with beans was put in front of their strong hand. After that participants were allowed to move buckets to their convenience as long as they did not violate the basic setup.

Then the task was started. Exercising the real-effort task participants picked beans with their right (left) hand and collected beans in their left (right) hand and emptied their left (right) hand into the red bucket. The enumerators stood inside the circle making sure that participants do not turn their heads to watch each other. After the time was up the enumerators collected the red buckets from the participants and weighed their content on a scale. This was our measure for real-effort output. The achieved weight in grams was recorded on the participant's card. Payoffs were determined according to the contract type and the achieved weight (except for the wage contract were the payoff was fixed). Payments were made at the end of the data collection.

4 Identification and results

4.1 Identification and inference

Given the session-wise randomization of the contract treatment (rather than randomization at the individual level) there are some identification issues I have to deal with. In total, there are 25 sessions. Sharecropping, fixed-rent and ownership treatment are respectively randomized across 6 sessions, the wage treatment across 7 sessions. However, given the fairly low number of randomizations per contract treatment, it is not straightforward to assume that the different treatments are uncorrelated with observables and unobservables.

To check for this, I regress household characteristics (table 3, denoted as controls A), production characteristics for own plots (table 4, controls B), production characteristics for sharecropped and rented-in plots (table 5, controls C), shocks affecting households (table 6, controls D) and government programs households have benefited from (table 7, controls E) on the different contract treatments (with the wage treatment being the reference treatment set to zero). See appendix B for a full description of all control variables employed in this study.

I find that the sharecropping treatment correlates with 3 out of 29 control variables (positive correlation: being literate, being affected by a health shock; negative correlation: using a hoe on own plots). The fixed-rent treatment correlates with 4 out of 29 controls (positive correlation: usage of fertilizer on own plots, usage of hoes on own plots; negative correlation: employing wage labor on sharecropped or rented-in plots and the existence of the government infrastructure programs) and the ownership treatment with 4 out of 29 controls (positive correlation: being affected by a health shock, being affected by an increase in prices for inputs; negative correlation: household size, having at least on plot under a fixed-rent arrangement).

While these correlations may simply be an α-error (given the null hypothesis that randomization worked), I cannot exclude a priori that the relatively higher occurrence of health shocks under the sharecropping and ownership treatment, for instance, affected effort significantly through a channel other than the contract treatment (because participants were impaired in exercising the real-effort task as a consequence of a health shock).

To check for this I estimate the regressions (weighted and unweighted) for the different samples (full sample, subsample of white and yellow beans, subsamples excluding sessions 7 and 13) controlling for the sets of control variables described above. Finding that the point estimates of the contract treatments are rather insensitive to the inclusion of a variety of control variables gives me confidence that unobserved factors may not affect real-effort output in a systematic way (despite the low number of randomizations per contract treatment).

Further, treatment assignment at the session level may cause conventional and robust standard errors to be biased. This is particularly difficult to deal with when the number of sessions is relatively small. Therefore I apply a procedure suggested by Donald and Lang (2007) and make inference on treatment effects

estimated with session means (see section 4.4 for more details).

4.2 Real-effort output and contract treatments (unweighted regressions)

I begin with results of the unweighted regressions. **Table 8** shows the estimated treatment effects for the full sample (that is, including white and yellow beans). Across all specifications (that is, successively including no controls, controls A, controls B, controls C, controls D, controls E and controls A to E altogether into the regression) the effect of the sharecropping treatment is negative. The sign of the treatment effect is opposite to what I expected but not statistically significant from the wage treatment. In turn, across all specifications the fixed-rent treatment is highly significant at the 1 percent level (inference in this and the following section is based on robust standard errors. Section 4.4 considers the possibility of clustering effects). As compared to the wage treatment, the fixed-rent treatment increases real-effort output between 8.9 and 9.6 percent. The effect of the ownership treatment is also significant, but smaller in magnitude. It is estimated to increase real-effort output between 6.4 and 7.2 percent.

As detailed in the experimental procedure, white and yellow beans are similar in size but slightly differ in weight. Therefore I repeat the estimation for both subsamples of white and yellow beans.

Table 9 presents the results for the white beans subsample. The overall picture is not as clear as for the full sample, but the most important findings are confirmed. The fixed-rent treatment increases output between 6.7 and 9.1 percent as compared to the wage treatment. All coefficients are highly significant. The ownership treatment increases output vis-a-vis the wage treatment between 2.8 and 6.0 percent. However, only 4 out of 6 specifications are found to be significant. Point estimates from regressions including household characteristics (controls A) almost half for the subsample as compared to the full sample and turn insignificant. The direction of the effect of the sharecropping treatment is again unexpected. For this subsample it even turns negative and statistically significant. As compared to the wage treatment it is estimated to decrease output between 5.7 and 7.4 percent.

However, looking at the subsample with yellow beans in **table 10** strongly supports the overall findings from the full sample in table 8. Compared to the wage treatment, the fixed-rent treatment increases output between 13.1 and 14.9 percent and the owner treatment between 11.0 and 13.4 percent. All of these coefficients are highly significant. For this subsample, the sharecropping treatment increases output between 3.9 and 5.5 percent (not significant).

Although the enumerators and I attempted to make the experimental data collection as comparable as possible, collecting data in the field never allows for a perfectly controlled environment as in lab. Two sessions differed notably in surrounding conditions.

In session 7 our experimental site happened to be neighbored by a market and an alternative site could not be organized due to time constraints. As a consequence this was the only session with an audience. I cannot exclude that

this had an effect on the exercised effort of participants, i.e. that participants were more motivated to exercise high effort due to spectators.

In session 13, the last of 7 sessions in the Woreda Gozamen, we got to know during the conversations with the Kebele chief, that some of the participants had heard about the experiment from relatives in a Kebele where we collected data three days before. Conversations with the Kebele chief and participants before and after the experiment were obligatory in all sessions. There were no signs that this problem occurred in other sessions as well.

To check the robustness of the results with respect to these sessions I excluded the two sessions from the sample and rerun the regressions for the full sample as well as the subsample of white and yellow beans. All main findings are robust to excluding these two sessions.

Table 11 shows the results for the full sample minus session 7 and 13. As compared to the wage treatment, the fixed-rent treatment increases output between 10.1 and 11.7 percent and the owner treatment between 7.6 and 9.4 percent.

4.3 Real-effort output and contract treatments (weighted regressions)

The random sample at hand raises the question as to whether these results are representative for the population of interest (that is, households in the Amhara and Oromia region). The results in the preceding section are based on a random sample, but do not consider unequal selection probabilities of households into the sample. Hence I attached weights to the regressions to produce point estimates as if each household in the sample represents an equal number of households in the population and rerun the regressions from the preceding section.

Table 12 shows the results for the full sample. Real-effort output increases between 12.2 and 13.1 percent for the fixed-rent treatment (all significant). Point estimates for the owner treatment range from 4.5 to 4.8 percent (significant in 3 out of 6 regressions) and between -2.8 and -2.2 percent for the sharecropping treatment (all insignificant). As compared to the unweighted regressions in table 8, point estimates for the fixed-rent treatment inflate. They slightly deflate for the owner and sharecropping treatment (implying that the unweighted regressions actually underestimate the effect of the fixed-rent treatment and overestimate the effect of the other two).

Attaching weights to the regressions based on the subsample with white beans (see **table 13**) leaves point estimates for the sharecropping and fixed-rent treatment nearly unchanged as compared to table 9. Both are significant across all specifications. However, the point estimates for the owner treatment become smaller and turn insignificant.

Table 14 presents results for weighted regressions based on the subsample with yellow beans. Estimated coefficients on fixed-rent and owner treatment are highly significant across all specifications, the sharecropping treatment is insignificant. As compared to table 10, point estimates on the fixed-rent treatment inflate through the inclusion of weights.

For reasons detailed above, I exclude session 7 and 13 for the weighted regressions as well. **Table 15** presents the results. Point estimates slightly inflate for the fixed-rent treatment (they range from 13.0 to 13.9 percent, all significant) and deflate for the owner treatment (they range from 4.8 to 5.7 percent, all significant). Estimated coefficients for the sharecropping treatment are nearly the same for weighted and unweighted regressions.

4.4 Clustering effects

Contract treatments vary only at the session level. Households within these sessions may be similar to each other because they come from the same Woreda and share similar characteristics (e.g. ethnicity, weather conditions). This implies a positive within session correlation. Consequently, conventional (and robust) standard errors may be biased.

Clustered standard errors as in Liang and Zeger (1986) are no solution to the problem at hand because consistency depends on the number of sessions getting large rather than session size. The number of sessions in my case is $j = 25$ and rather small. Therefore I follow a procedure by Donald and Lang (2007) who suggest using session means in estimation to account for clustering effects when treatment varies at the session level and the number of sessions is small.

Doing so, I regress log real-effort output on a full set of session dummies and various sets of control variables (same specification as in regressions above) at the first-stage. The session dummies represent covariate-adjusted session effects. At the second stage session dummies are regressed on the contract treatments (reducing the sample size to the number of sessions: $j = 25$ for the full sample and $j = 13$ and $j = 12$ for the subsample with white and yellow beans). Table 16 to 19 present the results. I follow the suggestion by Angrist and Pischke (2009: p. 321) and present both p-values based on conventional and robust standard errors (and use the larger of which for inference) as heteroscedasticity in the clustered residuals may bias robust standard errors.

Table 16 shows the results for the full sample. Point estimates are identical for the case without control variables (column 1) and quite similar for cases with control variables (other columns) as compared to table 8. At conventional levels (i.e. at the 10 percent level) fixed-rent and owner treatment cannot be inferred to differ significantly from the wage treatment for the full sample, However, p-values are close to the conventional rejection level. This result is mainly driven by the subsample of white beans (**table 17**). Looking at the subsample of yellow beans (**table 18**) actually confirms the inference on the findings in the preceding sections. The fixed-rent and owner treatment can be inferred to differ from the wage treatment at least at the 5 percent level. However, for the full sample outcomes are dominated by the subsample with white beans because they produce higher real-effort on average across all sessions (because white beans weight slightly more as compared to yellow beans).

Excluding session 7 and 13 from the sample increases the precision of the estimates and therefore results in lower p-values (see **tables 19**).

4.5 Order of experiments

Data collection in Kebeles consisted of two experiments (the real effort experiment and a risk attitude game similar to Tanaka et al., 2010) and a qualitative and quantitative interview. The interview was conducted last during all data collections. However, the order of the two experiments was randomized across all data collection sessions to avoid a systematic spillover from one experiment into the other. For instance, when the risk attitude game was conducted first, participants knew that they earned (or lost) some money already. Although payouts only took place at the end of the experiment, this may have altered the work effort of participants and therefore may affect outcomes through channels other than the treatment. Table 20 to 21 show the results of regressing log real-effort output on a binary indicator equal to one if the risk experiment was conducted first and zero if the real-effort experiment was conducted second for the unweighted sample (**table 20**) and the weighted sample (**table 21**). It is shown that the order of the game affects log real effort output in none of the regressions. I conclude that there was no systematic spillover from one experiment into the other.

5 Is there a real life spillover into the experiment?

Factors which rationally determine agricultural contract choice and hence input intensity and output in real life should not matter for effort provision in an experimental setting where contract treatment assignment is independent of these factors. However, if the outcomes of the experiment are (at least partially) unexpected, it raises the question whether there is a real life spillover into the experiment affecting the behavior of participants (at least for a subgroup of participants).

The experiment was accompanied by semi-structured interviews on agricultural contracts which might help shed some light on why farmers do not differ in their response to wage and sharecropping treatment. Across almost all study sites participants expressed discontent with sharecropping contracts. To quote a typical answer:

> Sharecropping is not my first preference, but the land owners forces this land arrangement upon me. Because I only have limited plot size I have to accept sharecropping.
>
> (A Farmer interviewed in the Kebele Yetejan)

This raises the question why farmers go into land markets in the first place. Rural areas in Amhara and Oromia are densely populated and fertile land is scarce. Big family sizes lead to the defragmentation of plots as land has to be shared among many inheritors. Further, a highly controversial land reform took place in Amhara in 1997 (but not in Oromia). For political reasons, it has

favored some groups, deprived others and produced in its aftermath a substantial number of youngsters pushed into land markets (Ege, 1997).

Thus, asking farmers why they go into land markets a typical answer is:

> Because the land I owe is not enough.
> (A Farmer interviewed in the Kebele Emesa Spsoto)

Besides that a great deal of discontent stems from the terms of the sharecropping contracts. This is a point made in Deininger et al. (2011), who find that sharecropped plots yield 16 to 25 percent less output than owned plots cultivated by the same households. The inefficiency disappears for sharecropping contracts where input costs are shared. This, however, is very uncommon in Ethiopia (in the sample of Deininger et al. only 12.1 percent of all sharecropping contracts have a cost-sharing component). With respect to typical contracts farmers gave statements as follows:

> Sharecroping is the better arrangement for the landowner. He gets half of the output without contributing work or other inputs to the crop production process.
> (A Farmer interviewed in the Kebele Addisena Guilt)

Given the apparent discontent with sharecropping contracts, I explored whether this may have spilled over into the experiment. I check whether real life sharecroppers assigned to the sharecropping treatment (roughly 10.6 percent of the sample) exercise real-effort different from others assigned to the sharecropping treatment. Doing so, I regress real-effort output on the contract treatments (with the wage treatments being the reference), a binary dummy equal to one if a participant is a real life sharecropper and zero if not and an interaction term of this binary indicator with the sharecropping treatment.

Table 22 shows the results. Across all specifications the main effect of being a real life sharecropper does not have a significant effect on real-effort output. However, the interaction term is negative and significant. This implies that real life sharecroppers exercise significantly less effort than non-sharecroppers in the same treatment.

In **table 23** I undertake a similar analysis, checking whether real life sharecroppers assigned to the fixed-rent treatment (roughly 9.6 percent of the sample). Thus I regress real-effort output on the contract treatments (with the wage treatments being the reference), a binary dummy equal to one if a participant is a real life sharecropper and zero if not and an interaction term of this binary indicator with the fixed-rent treatment.

Table 24 and **table 25** repeat exercises with weighted regressions and largely confirm the results from table 23 and 24. These results imply that discontent with real life sharecropping contracts may impair the pure incentive effect of the sharecropping treatment in the experiment.

6 Concluding remarks

The findings demonstrate clear productivity effects of fixed-rent and owner treatment as compared to wage treatment. The sharecropping and wage treatment cannot be distinguished in productivity effects (with the exception of one subsample where the sharecropping treatment actually yields lower productivity as the wage treatment).

With respect to fixed-rent and ownership contracts this confirms basic economic theory. These contracts induce the tenant to supply effort up to the point where the marginal product of the effort equals their marginal cost. Effort exercised under the sharecropping treatment should be lower as tenants receive only a fraction of their marginal product. However, given that wage payments are sufficiently low, all of these contracts should induce higher effort according to basic economic theory. Our experimental results are clearly in line with theory for fixed-rent and owner treatment as compared to the wage treatment, but clearly not for the sharecropping treatment. This comes as a surprise.

Drawing on semi-structured interviews farmers express their discontent with sharecropping contracts at almost all study sites, which seems to spillover into the experiment. Real life sharecroppers randomly assigned to the sharecropping treatment clearly exercise less effort than non-sharecroppers, while they exercise more effort under the fixed-rent treatment. The incentive provision of the sharecropping treatment seems to be impaired by the discontent with the sharecropping contracts in real life.

I do not want to explore farmers actual motives for sharecropping here (this I do elsewhere). I also do not want to overstretch the external claims that follow from this experiment, but given the apparent discontent with widespread sharecropping contracts this deserves further attention in future research.

References

[1] Abeler, J., Falk, A., Goette, L. and D. Huffman (2011) "Reference Points and Effort Provision," *American Economic Review*, 101(2): 470-92.

[2] Ackerberg, D. and M. Botticini (2002) "Endogenous Matching and the Empirical Determinants of Contract Form," *Journal of Political Economy*, 110(3): 564-91.

[3] D'Andrea, A., Lyons, D., Haile, M. and E. Butler (1999) "Ethnoarchaeological Approaches to the Study of Prehistoric Agriculture in the Ethiopian Highlands," in V. der Veen (Ed.) *The Exploitation of Plant Resources in Ancient Africa*, Kluwer Academic/Plenum Publishers: New York.

[4] Angrist, J. and J.-S. Pischke (2009) *Mostly Harmless Econometrics*, Princeton University Press: Princeton and Oxford.

[5] Cheung, S. (1969) *The Theory of Share Tenancy*, University of Chicago Press: Chicago.

[6] Deaton, A. (1997) *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*, Johns Hopkins University Press: Baltimore, Maryland.

[7] Deininger, K., Ali, A., D. and T. Alemu (2008) "Assessing the Functioning of Land Rental Markets in Ethiopia," *Economic Development and Cultural Change*, 57(1), 67-100.

[8] Deininger, K., Ali, A. and T. Alemu (2011) "Productivity effects of land rental markets in Ethiopia : Evidence from a matched tenant-landlord sample," *World Bank Policy Research Working Paper Series*, 5727.

[9] de Wet, J. (1977) "Domestication of African Cereals," *African Economic History*, 3, 15-32.

[10] Dohmen, T. and A. Falk (2011) "Performance Pay and Multidimensional Sorting: Productivity, Preferences, and Gender," *American Economic Review*, 101, 556–90.

[11] Donald, S. and K. Lang (2007) "Inference with Difference-in-Difference and other Panel Data," *Review of Economics and Statistics*, 89, 221-33.

[12] Ege, S. (1997) The Promised Land: The Amhara Land Redistribution of 1997, SMU-rapport 5/97, Norwegian University of Science and Technology, Centre for Environment and Development, Dragvoll, Norway.

[13] Ehret, C. (1979) "On the Antiquity of Agriculture in Ethiopia," *The Journal of African History*, 20(2), 161-77.

[14] Ghebru, H. and S. Holden (2006) "Factor Market Imperfections and Rural Markets in Northern Ethiopian Highlands."

[15] Hallagan, W. (1978) "Self-selection by Contractual Choice and the Theory of Sharecropping," *Bell Journal of Economics*, 9(2), 344-54.

[16] Hamilton, B., Nickerson, J. and H. Owan (2003) "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation," *Journal of Political Economy*, 111(3), 465– 97.

[17] Holden, S. and H. Ghebru (2006) "Kinship, Transaction Costs and Land Rental Market Participation."

[18] Lazear, E. (1989) "Pay Equality and Industrial Politics," *Journal of Political Economy*, 97(3), 1261–84.

[19] Lazear, E. (2000) "Performance Pay and Productivity," *American Economic Review*, 90(5), 1346 –61.

[20] Liang, K. and S. Zeger (1986) "Longitudinal Data Analysis using Generalized Linear Models," *Biometrika*, 73, 13-22.

[21] Newbery, D. and J. Stiglitz (1979) "Sharecropping, Risk-sharing, and the importance of Imperfect Information." In J. Roumasset, J. Boussard and I. Singh (Eds.), *Risk, Uncertainty and Agricultural Development*, Ch. 17. New York: Agricultural Development Council.

[22] Otsuka, K., Chuma, H. and Y. Hayami (1992) "Land and Labor Contracts in Agrarian Economies: Theories and Facts," *Journal of Economic Literature*, 30(4), 1965-2018.

[23] Pender, J. and M. Fafchamps (2006) "Land Lease Markets and Agricultural Efficiency in Ethiopia," *Journal of African Economies*, 15(2), 251-84.

[24] Shearer, B. (2004) "Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment," *Review of Economic Studies*, 71, 513–34.

[25] Stiglitz, J. (1974) "Incentives and Risk Sharing in Sharecropping," *Review of Economic Studies*, 41(2), 219-55.

[26] Tanaka, T., Camerer, C. and Q. Nguyen (2010) "Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam," *American Economic Review*, 100(1), 557–71.

[27] Teklu, T. and A. Lemi (2004) "Factors Affecting Entry and Intensity in Informal Rental Land Markets in Southern Ethiopian Highlands," *Agricultural Economics*, 30 (2), 117-28.

A Tables

Table 1: Correlations between real contracts and per-capita expenditure level.

	(1) PCEXPENDITURE	(2) PCEXPENDITURE	(3) PCEXPENDITURE	(4) PCEXPENDITURE
WAGE	23.6402			
	(58.7443)			
SHARECROP		-96.7808*		
		(50.9240)		
RENT			126.7766**	
			(56.2886)	
OWNER				135.7312**
				(67.5128)
Mean (s.d.) dependent variable	727.48 (571.85)	727.48 (571.85)	727.48 (571.85)	727.48 (571.85)
Mean (s.d.) control variable	0.8840 (0.3205)	0.4549 (0.4984)	0.2104 (0.4080)	0.9498 (0.2183)
Observations	498	498	498	498
R-squared	0.0002	0.0071	0.0081	0.0027

Notes: Bivariate ordinary least squares (OLS) regressions with robust standard errors in parentheses. Dependent variable is the per-capita expenditure level of households (PCEXPENDITURE) denominated in Ethiopian Birr. Following Deaton (1997) children between 0 and 5 years of age received the weight of 0.4 adults, children between 6 and 11 receive the weight of 0.5 adults. All household members older than 11 are considered adults and receive a weight of 1 in the calculation of per-capita expenditures. Wage (WAGE), sharecropping (SHARECROP), fixed-rent (RENT) and ownership (OWNER) contract are binary indicators equal to one if the respective contract was used by a household on at least one plot during the last cropping season and zero otherwise (see appendix B for a full description of variables). Significance level at 90(*), 95(**), 99(***) percent confidence. Differences between maximum sample size n=500 and observations are due to missing values.

Table 2: Contract treatments and real-effort output.

	(1) Number of sessions	(2) Number of participants	(3) Sample means (s.d.) of output	(4) Average payoff
C1: Wage contract	**7**	**140**	**137.20 (29.62)**	**20.00 (0.00)**
Only white beans	4	80	150.61 (26.47)	20.00 (0.00)
Only yellow beans	3	60	119.32 (23.66)	20.00 (0.00)
C2: Sharecroping contract	**6**	**120**	**132.14 (23.10)**	**26.43 (4.62)**
Only white beans	3	60	139.92 (21.04)	27.98 (4.21)
Only yellow beans	3	60	124.37 (22.59)	24.87 (4.52)
C3: Fixed-rent contract	**6**	**120**	**149.44 (24.55)**	**39.78 (9.83)**
Only white beans	3	60	161.88 (20.92)	44.75 (8.37)
Only yellow beans	3	60	137.00 (21.53)	34.80 (8.61)
C4: Ownership contract	**6**	**120**	**146.08 (26.14)**	**58.43 (10.46)**
Only white beans	3	60	159.5166 (24.54)	63.80 (9.81)
Only yellow beans	3	60	132.63 (20.22)	53.05 (8.08)
Total	**25**	**500**	**141.05 (26.92)**	**35.51 (16.50)**
Only white beans	13	260	152.80 (24.92)	37.66 (18.17)
Only yellow beans	12	240	128.33 (22.97)	33.18 (14.15)

Notes: Descriptive statistics on number of sessions, participants, exercised real-effort output and average payoff of participants for the different contract treatments in the experiment. The wage treatment was randomly assigned to 7 sessions, sharecropping, fixed-rent and ownership contracts to respectively 6 sessions. Each session had 20 participants (25 sessions with 500 participants in total). Each treatment was played 3 times with white beans and 3 times with yellow beans, except for the wage treatment which was played 4 times with white beans (in total 13 sessions with white beans and 12 sessions with yellow beans). The color of beans were randomly assigned to treatments. Real-effort output is measured in grams. Payoff was determined by exercised effort and contract treatment except for the wage treatment where payoff was fixed. Payoffs were made in Ethiopian Birr. 20 Birr amounted roughly to 1.20 USD at the time of data collection.

Table 3: Contract treatments and household characteristics (Controls A).

	(1) GENDER	(2) AGE	(3) AGESQ	(4) MEMBER	(5) LITERATE	(6) HHSIZE
C2	-0.0352	0.6385	51.6692	-0.0005	0.1449**	0.2238
	(0.0328)	(1.3253)	(123.6850)	(0.0553)	(0.0599)	(0.2654)
C3	0.0481	2.1052	180.4359	-0.0005	-0.0846	0.0155
	(0.0405)	(1.3245)	(122.0387)	(0.0553)	(0.0625)	(0.2533)
C4	0.0241	-1.2166	-91.8660	-0.0279	-0.0089	-0.4555*
	(0.0387)	(1.3706)	(124.7409)	(0.0564)	(0.0625)	(0.2377)
Mean (s.d.) dependend variable	0.1004 (0.3008)	43.08 (10.75)	1971.42 (995.05)	0.7269 (0.4460)	0.5593 (0.4969)	5.89 (1.99)
Observations	498	498	498	498	497	499
R-squared	0.0100	0.0119	0.0094	0.0007	0.0270	0.0148

Notes: Ordinary least squares (OLS) regressions with robust standard errors in parentheses. Dependent variables are household characteristics such as head's gender (GENDER), head's age (AGE), head's age squared (AGESQ), head's membership status in a socio-political organization (MEMBER), head's literacy status (LITERATE) and household site (HHSIZE). Treatment variables are sharecropping contract (C2), fixed-rent contract (C3) and ownership contract (C4), all binary indicators equal to one if the respective contract treatment was randomly assigned to a session and zero otherwise. The wage treatment (C1) is the reference contract (see appendix B for a full description of variables). Significance level at 90(*), 95(**), 99(***) percent confidence. If randomization of contracts was perfectly successful, household characteristics should be uncorrelated with contract treatments. Differences between maximum sample size n=500 and observations are due to missing values.

Table 4: Contract treatments and production characteristics own plots (Controls B).

	(1) WAGE	(2) FERTILIZER	(3) HOE	(4) PLOUGH	(5) SADDLE	(6) THRESHING
C2	0.0095	0.0405	-0.0798*	0.0917	0.0524	-0.0488
	(0.0430)	(0.0348)	(0.0415)	(0.0622)	(0.0609)	(0.0339)
C3	0.0595	0.0571*	0.1036**	-0.0167	0.0440	0.0095
	(0.0390)	(0.0330)	(0.0519)	(0.0620)	(0.0608)	(0.0394)
C4	0.0429	-0.0189	0.0702	-0.0167	-0.0143	0.0179
	(0.0405)	(0.0403)	(0.0506)	(0.0620)	(0.0598)	(0.0401)
Mean (s.d.) dependend variable	0.8840 (0.3205)	0.9118 (0.2838)	0.1940 (0.3958)	0.464 (0.4992)	0.3840 (0.4868)	0.1020 (0.3029)
Observations	500	499	500	500	500	500
R-squared	0.0058	0.0113	0.0305	0.0079	0.0033	0.0070

Notes: Ordinary least squares (OLS) regressions with robust standard errors in parentheses. Dependent variables are production characteristics such as whether households have employed wage labor (WAGE) or used fertilizer (FERTILIZER), a hoe (HOE), a plough (PLOUGH), a saddle (SADDLE) or a threshing machine (THRESHING) on the plots they own during the last year. Treatment variables are sharecropping contract (C2), fixed-rent contract (C3) and ownership contract (C4), all binary indicators equal to one if the respective contract treatment was randomly assigned to a session and zero otherwise. The wage treatment (C1) is the reference contract (see appendix B for a full description of variables). Significance level at 90(*), 95(**), 99(***) percent confidence. If randomization of contracts was perfectly successful, production characteristics should be uncorrelated with contract treatments. Differences between maximum sample size n=500 and observations are due to missing values.

Table 5: Contract treatments and production characteristics for inputs on sharecropped and rented-in plots (Controls B cont.).

	(1) SHARECROP	(2) RENT	(3) FERTILIZER	(4) WAGE	(5) PLOUGH	(6) SADDLE
C2	-0.0226	0.0560	0.0988	0.0214	0.0821	0.0310
	(0.0621)	(0.0551)	(0.0619)	(0.0533)	(0.0588)	(0.0542)
C3	-0.0643	-0.0274	0.0321	-0.0869*	-0.0595	-0.0024
	(0.0617)	(0.0518)	(0.0624)	(0.0479)	(0.0547)	(0.0529)
C4	0.0483	-0.1349***	-0.0223	-0.0536	-0.0012	-0.0690
	(0.0625)	(0.0454)	(0.0625)	(0.0498)	(0.0568)	(0.0496)
Mean (s.d.) dependend variable	0.4549 (0.4984)	0.2104 (0.4080)	0.5190 (0.5001)	0.2000 (0.4004)	0.2980 (0.4578)	0.2260 (0.4186)
Observations	499	499	499	500	500	500
R-squared	0.0064	0.0279	0.0081	0.0112	0.0117	0.0073

Notes: Ordinary least squares (OLS) regressions with robust standard errors in parentheses. Dependent variables are production characteristics such as whether households sharecropped (SHARECROP) or rented-in (RENT) a plot and used fertilizer (FERTILIZER), employed wage labor (WAGE), a plough (PLOUGH) or a saddle (SADDLE) on plots they sharecropped or rented-in. Treatment variables are sharecropping contract (C2), fixed-rent contract (C3) and ownership contract (C4), all binary indicators equal to one if the respective contract treatment was randomly assigned to a session and zero otherwise. The wage treatment (C1) is the reference contract (see appendix B for a full description of variables). Significance level at 90(*), 95(**), 99(***) percent confidence. If randomization of contracts was perfectly successful, production characteristics should be uncorrelated with contract treatments. Differences between maximum sample size n=500 and observations are due to missing values.

Table 6: Contract treatments and shocks (Controls C)

	(1) DROUGHT	(2) PEST	(3) LDEATH	(4) DROPSALE	(5) PRICEFOOD	(6) HEALTH	(7) PRICEINPUT
C2	-0.0119	-0.0238	0.0631	0.0762	-0.0107	0.0762*	-0.0857
	(0.0471)	(0.0536)	(0.0509)	(0.0537)	(0.0583)	(0.0441)	(0.0618)
C3	0.0214	0.0429	0.0298	0.0345	0.0226	0.0429	0.0893
	(0.0490)	(0.0560)	(0.0494)	(0.0522)	(0.0573)	(0.0420)	(0.0621)
C4	0.0048	0.0179	-0.0369	0.0595	-0.0440	0.1179**	0.1560**
	(0.0481)	(0.0552)	(0.0456)	(0.0532)	(0.0590)	(0.0464)	(0.0611)
Mean (s.d.) dependend	0.1820	0.266	0.192	0.248	0.678	0.1640	0.5240
variable	(0.3862)	(0.4423)	(0.3943)	(0.4322)	(0.4677)	(0.3706)	(0.4999)
Observations	500	500	500	500	500	500	500
R-squared	0.0009	0.0029	0.0085	0.0046	0.0025	0.0141	0.0323

Notes: Ordinary least squares (OLS) regressions with robust standard errors in parentheses. Dependent variables are shock characteristics such as whether households experienced a drought (DROUGHT), crop pest (PEST), livestock death (LDEATH), a sharp drop in sales prices for their agricultural produce (DROPSALES), a sharp increase in the price of food they are buying elsewhere (PRICEFOOD), a household member having a health shock (HEALTH) and households being affected by a strong increase in the price for agricultural inputs (PRICEINPUT) during the last year. Treatment variables are sharecropping contract (C2), fixed-rent contract (C3) and ownership contract (C4), all binary indicators equal to one if the respective contract treatment was randomly assigned to a session and zero otherwise. The wage treatment (C1) is the reference contract (see appendix B for a full description of variables). Significance level at 90(*), 95(**), 99(***) percent confidence. If randomization of contracts was perfectly successful, shocks should be uncorrelated with contract treatments. Differences between maximum sample size n=500 and observations are due to missing values.

Table 7: Contract treatments and government programs (Controls D)

	(1) SLM	(2) INFRASTRUCTURE	(3) SAVINGS	(4) SELFHELP
C2	-0.0560	-0.0857	-0.0464	0.0048
	(0.0551)	(0.0589)	(0.0620)	(0.0562)
C3	-0.0393	-0.1274**	0.0536	-0.0119
	(0.0545)	(0.0576)	(0.0624)	(0.0556)
C4	-0.0226	-0.0524	0.0952	-0.0286
	(0.0539)	(0.0598)	(0.0621)	(0.0550)
Mean (s.d.) dependend variable	0.7360 (0.4412)	0.3220 (0.4677)	0.4960 (0.5005)	0.2700 (0.4444)
Observations	500	500	500	500
R-squared	0.0023	0.0103	0.0111	0.0008

Notes: Ordinary least squares (OLS) regressions with robust standard errors in parentheses. Dependent variables are indicators stating whether households have benefited from government programs such as the sustainable land managment (SLM), infrastructure (INFRASTRUCTURE), irrigation expension (IRRIGATION), savings and credit (SAVINGS) and self-help group (SELFHELP) program in the past. Treatment variables are sharecropping contract (C2), fixed-rent contract (C3) and ownership contract (C4), all binary indicators equal to one if the respective contract treatment was randomly assigned to a session and zero otherwise. The wage treatment (C1) is the reference contract (see appendix B for a full description of variables). Significance level at 90(*), 95(**), 99(***) percent confidence. If randomization of contracts was perfectly successful, government programs should be uncorrelated with contract treatments. Differences between maximum sample size n=500 and observations are due to missing values.

Table 8: Contract treatments and real-effort output (white and yellow beans, unweighted regressions)

	(1) EFFORT	(2) EFFORT	(3) EFFORT	(4) EFFORT	(5) EFFORT	(6) EFFORT
C2	-0.0288	-0.0240	-0.0307	-0.0301	-0.0291	-0.0298
	(0.0254)	(0.0236)	(0.0269)	(0.0256)	(0.0253)	(0.0247)
C3	0.0964***	0.0955***	0.0927***	0.0960***	0.0932***	0.0886***
	(0.0246)	(0.0234)	(0.0264)	(0.0244)	(0.0248)	(0.0250)
C4	0.0715***	0.0635***	0.0716***	0.0711***	0.0686***	0.0611***
	(0.0252)	(0.0231)	(0.0260)	(0.0249)	(0.0253)	(0.0236)
Controls A	no	yes	no	no	no	yes
Controls B	no	no	yes	no	no	yes
Controls C	no	no	no	yes	no	yes
Controls D	no	no	no	no	yes	yes
Mean (s.d.) dependend	4.93	4.93	4.93	4.93	4.93	4.93
variable	(0.1977)	(0.1977)	(0.1977)	(0.1977)	(0.1977)	(0.1977)
Observations	500	497	499	500	500	497
R-squared	0.0651	0.2211	0.0720	0.0777	0.0717	0.2410

Notes: Ordinary least squares (OLS) regressions with robust standard errors in parentheses. Dependent variable is log real-effort output (EFFORT) during the experiment (in levels measured in grams). Treatment variables are sharecropping contract (C2), fixed-rent contract (C3) and ownership contract (C4), all binary indicators equal to one if the respective contract treatment was randomly assigned to a session and zero otherwise. The wage treatment (C1) is the reference contract. Significance level at 90(*), 95(**), 99(***) percent confidence. Controls A include household characteristics, controls B include production characteristics of cultivated plots, controls C include household shock indicators and controls D include indicators on government programs (see appendix B for a full description of variables). Differences between maximum sample size n=500 and observations are due to missing values.

Table 9: Contract treatments and real-effort output (only white beans, unweighted regressions)

	(1) EFFORT	(2) EFFORT	(3) EFFORT	(4) EFFORT	(5) EFFORT	(6) EFFORT
C2	-0.0701**	-0.0569**	-0.0741**	-0.0606**	-0.0698**	-0.0580*
	(0.0277)	(0.0262)	(0.0302)	(0.0300)	(0.0289)	(0.0333)
C3	0.0786***	0.0912***	0.0672**	0.0829***	0.0763***	0.0843***
	(0.0259)	(0.0240)	(0.0268)	(0.0265)	(0.0272)	(0.0274)
C4	0.0601**	0.0315	0.0556*	0.0568*	0.0565*	0.0279
	(0.0283)	(0.0254)	(0.0283)	(0.0292)	(0.0297)	(0.0276)
Controls A	no	yes	no	no	no	yes
Controls B	no	no	yes	no	no	yes
Controls C	no	no	no	yes	no	yes
Controls D	no	no	no	no	yes	yes
Mean (s.d.) dependend	4.93	4.93	4.93	4.93	4.93	4.93
variable	(0.1977)	(0.1977)	(0.1977)	(0.1977)	(0.1977)	(0.1977)
Observations	260	260	260	260	260	260
R-squared	0.1146	0.3270	0.1580	0.1231	0.1213	0.3468

Notes: Ordinary least squares (OLS) regressions with robust standard errors in parentheses. Dependent variable is log real-effort output (EFFORT) during the experiment (in levels measured in grams). Treatment variables are sharecropping contract (C2), fixed-rent contract (C3) and ownership contract (C4), all binary indicators equal to one if the respective contract treatment was randomly assigned to a session and zero otherwise. The wage treatment (C1) is the reference contract. Significance level at 90(*), 95(**), 99(***) percent confidence. Controls A include household characteristics, controls B include production characteristics of cultivated plots, controls C include household shock indicators and controls D include indicators on government programs (see appendix B for a full description of variables). Differences between maximum subsample size n=260 and observations are due to missing values.

Table 10: Contract treatments and real-effort output (only yellow beans, unweighted regressions)

	(1) EFFORT	(2) EFFORT	(3) EFFORT	(4) EFFORT	(5) EFFORT	(6) EFFORT
C2	0.0469	0.0385	0.0550	0.0446	0.0461	0.0437
	(0.0374)	(0.0344)	(0.0396)	(0.0350)	(0.0380)	(0.0343)
C3	0.1486***	0.1310***	0.1602***	0.1489***	0.1486***	0.1398***
	(0.0349)	(0.0333)	(0.0369)	(0.0339)	(0.0347)	(0.0347)
C4	0.1172***	0.1337***	0.1246***	0.1104***	0.1180***	0.1316***
	(0.0343)	(0.0343)	(0.0338)	(0.0320)	(0.0341)	(0.0319)
Controls A	no	yes	no	no	no	yes
Controls B	no	no	yes	no	no	yes
Controls C	no	no	no	yes	no	yes
Controls D	no	no	no	no	yes	yes
Mean (s.d.) dependend	4.93	4.93	4.93	4.93	4.93	4.93
variable	(0.1977)	(0.1977)	(0.1977)	(0.1977)	(0.1977)	(0.1977)
Observations	240	237	239	240	240	237
R-squared	0.0965	0.2499	0.1394	0.1345	0.0981	0.3252

Notes: Ordinary least squares (OLS) regressions with robust standard errors in parentheses. Dependent variable is log real-effort output (EFFORT) during the experiment (in levels measured in grams). Treatment variables are sharecropping contract (C2), fixed-rent contract (C3) and ownership contract (C4), all binary indicators equal to one if the respective contract treatment was randomly assigned to a session and zero otherwise. The wage treatment (C1) is the reference contract. Significance level at 90(*), 95(**), 99(***) percent confidence. Controls A include household characteristics, controls B include production characteristics of cultivated plots, controls C include household shock indicators and controls D include indicators on government programs (see appendix B for a full description of variables). Differences between maximum subsample size n=240 and observations are due to missing values.

Table 11: Contract treatments and real-effort output excluding session 7 and 13 (white and yellow beans, unweighted regressions)

	(1) EFFORT	(2) EFFORT	(3) EFFORT	(4) EFFORT	(5) EFFORT	(6) EFFORT
C2	-0.0048	-0.0008	-0.0048	-0.0065	-0.0056	-0.0098
	(0.0281)	(0.0263)	(0.0301)	(0.0283)	(0.0280)	(0.0280)
C3	0.1140***	0.1095***	0.1124***	0.1167***	0.1114***	0.1058***
	(0.0262)	(0.0254)	(0.0279)	(0.0259)	(0.0264)	(0.0273)
C4	0.0891***	0.0771***	0.0880***	0.0938***	0.0865***	0.0763***
	(0.0268)	(0.0250)	(0.0277)	(0.0263)	(0.0269)	(0.0259)
Controls A	no	yes	no	no	no	yes
Controls B	no	no	yes	no	no	yes
Controls C	no	no	no	yes	no	yes
Controls D	no	no	no	no	yes	yes
Mean (s.d.) dependend variable	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)
Observations	460	457	459	460	460	457
R-squared	0.0701	0.2238	0.0789	0.0869	0.0759	0.2457

Notes: Ordinary least squares (OLS) regressions with robust standard errors in parentheses. Dependent variable is log real-effort output (EFFORT) during the experiment (in levels measured in grams). Treatment variables are sharecropping contract (C2), fixed-rent contract (C3) and ownership contract (C4), all binary indicators equal to one if the respective contract treatment was randomly assigned to a session and zero otherwise. The wage treatment (C1) is the reference contract. Significance level at 90(*), 95(**), 99(***) percent confidence. Controls A include household characteristics, controls B include production characteristics of cultivated plots, controls C include household shock indicators and controls D include indicators on government programs (see appendix B for a full description of variables). Differences between maximum subsample size n=460 and observations are due to missing values.

Table 12: Contract treatments and real-effort output (white and yellow beans, weighted regressions)

	(1) EFFORT	(2) EFFORT	(3) EFFORT	(4) EFFORT	(5) EFFORT	(6) EFFORT
C2	-0.0247	-0.0224	-0.0251	-0.0249	-0.0261	-0.0276
	(0.0270)	(0.0254)	(0.0286)	(0.0268)	(0.0270)	(0.0263)
C3	0.1304***	0.1278***	0.1303***	0.1307***	0.1277***	0.1224***
	(0.0263)	(0.0250)	(0.0280)	(0.0259)	(0.0264)	(0.0264)
C4	0.0455*	0.0450*	0.0467	0.0475*	0.0436	0.0424
	(0.0276)	(0.0259)	(0.0284)	(0.0267)	(0.0277)	(0.0264)
Controls A	no	yes	no	no	no	yes
Controls B	no	no	yes	no	no	yes
Controls C	no	no	no	yes	no	yes
Controls D	no	no	no	no	yes	yes
Mean (s.d.) dependend variable	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)
Observations	500	497	499	500	500	497
R-squared	0.0855	0.2305	0.0903	0.0987	0.0915	0.2483

Notes: Ordinary least squares (OLS) regressions with robust standard errors in parentheses. Dependent variable is log real-effort output (EFFORT) during the experiment (in levels measured in grams). Treatment variables are sharecropping contract (C2), fixed-rent contract (C3) and ownership contract (C4), all binary indicators equal to one if the respective contract treatment was randomly assigned to a session and zero otherwise. The wage treatment (C1) is the reference contract. Significance level at 90(*), 95(**), 99(***) percent confidence. Controls A include household characteristics, controls B include production characteristics of cultivated plots, controls C include household shock indicators and controls D include indicators on government programs (see appendix B for a full description of variables). Differences between maximum sample size n=500 and observations are due to missing values.

Table 13: Contract treatments and real-effort output (only white beans, weighted regressions)

	(1) EFFORT	(2) EFFORT	(3) EFFORT	(4) EFFORT	(5) EFFORT	(6) EFFORT
C2	-0.0838***	-0.0692**	-0.0947***	-0.0697**	-0.0850***	-0.0794**
	(0.0291)	(0.0271)	(0.0315)	(0.0309)	(0.0299)	(0.0325)
C3	0.0772***	0.0840***	0.0608**	0.0892***	0.0746***	0.0770***
	(0.0277)	(0.0259)	(0.0280)	(0.0278)	(0.0287)	(0.0278)
C4	0.0490	0.0197	0.0391	0.0527*	0.0457	0.0154
	(0.0305)	(0.0277)	(0.0306)	(0.0311)	(0.0318)	(0.0298)
Controls A	no	yes	no	no	no	yes
Controls B	no	no	yes	no	no	yes
Controls C	no	no	no	yes	no	yes
Controls D	no	no	no	no	yes	yes
Mean (s.d.) dependend	4.93	4.93	4.93	4.93	4.93	4.93
variable	(0.1977)	(0.1977)	(0.1977)	(0.1977)	(0.1977)	(0.1977)
Observations	260	260	260	260	260	260
R-squared	0.1385	0.3473	0.1823	0.1572	0.1482	0.3765

Notes: Ordinary least squares (OLS) regressions with robust standard errors in parentheses. Dependent variable is log real-effort output (EFFORT) during the experiment (in levels measured in grams). Treatment variables are sharecropping contract (C2), fixed-rent contract (C3) and ownership contract (C4), all binary indicators equal to one if the respective contract treatment was randomly assigned to a session and zero otherwise. The wage treatment (C1) is the reference contract. Significance level at 90(*), 95(**), 99(***) percent confidence. Controls A include household characteristics, controls B include production characteristics of cultivated plots, controls C include household shock indicators and controls D include indicators on government programs (see appendix B for a full description of variables). Differences between maximum subsample size n=260 and observations are due to missing values.

Table 14: Contract treatments and real-effort output (only yellow beans, weighted regressions)

	(1) EFFORT	(2) EFFORT	(3) EFFORT	(4) EFFORT	(5) EFFORT	(6) EFFORT
C2	0.0477	0.0369	0.0545	0.0443	0.0462	0.0407
	(0.0372)	(0.0346)	(0.0387)	(0.0348)	(0.0380)	(0.0347)
C3	0.1716***	0.1548***	0.1828***	0.1701***	0.1704***	0.1598***
	(0.0361)	(0.0345)	(0.0374)	(0.0351)	(0.0361)	(0.0357)
C4	0.1092***	0.1225***	0.1176***	0.1023***	0.1090***	0.1197***
	(0.0352)	(0.0352)	(0.0348)	(0.0332)	(0.0352)	(0.0333)
Controls A	no	yes	no	no	no	yes
Controls B	no	no	yes	no	no	yes
Controls C	no	no	no	yes	no	yes
Controls D	no	no	no	no	yes	yes
Mean (s.d.) dependend variable	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)
Observations	240	237	239	240	240	237
R-squared	0.1051	0.2511	0.1520	0.1435	0.1062	0.3233

Notes: Ordinary least squares (OLS) regressions with robust standard errors in parentheses. Dependent variable is log real-effort output (EFFORT) during the experiment (in levels measured in grams). Treatment variables are sharecropping contract (C2), fixed-rent contract (C3) and ownership contract (C4), all binary indicators equal to one if the respective contract treatment was randomly assigned to a session and zero otherwise. The wage treatment (C1) is the reference contract. Significance level at 90(*), 95(**), 99(***) percent confidence. Controls A include household characteristics, controls B include production characteristics of cultivated plots, controls C include household shock indicators and controls D include indicators on government programs (see appendix B for a full description of variables). Differences between maximum subsample size n=240 and observations are due to missing values.

Table 15: Contract treatments and real-effort output excluding session 7 and 13 (white and yellow beans, weighted regressions)

	(1) EFFORT	(2) EFFORT	(3) EFFORT	(4) EFFORT	(5) EFFORT	(6) EFFORT
c2	-0.0146	-0.0120	-0.0134	-0.0139	-0.0164	-0.0187
	(0.0289)	(0.0273)	(0.0309)	(0.0287)	(0.0288)	(0.0287)
c3	0.1373***	0.1330***	0.1382***	0.1393***	0.1348***	0.1295***
	(0.0272)	(0.0261)	(0.0290)	(0.0268)	(0.0274)	(0.0280)
c4	0.0524*	0.0498*	0.0532*	0.0572**	0.0506*	0.0483*
	(0.0285)	(0.0268)	(0.0294)	(0.0276)	(0.0286)	(0.0279)
Controls A	no	yes	no	no	no	yes
Controls B	no	no	yes	no	no	yes
Controls C	no	no	no	yes	no	yes
Controls D	no	no	no	no	yes	yes
Mean (s.d.) dependend variable	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)
Observations	460	457	459	460	460	457
R-squared	0.0867	0.2317	0.0932	0.1019	0.0916	0.2497

Notes: Ordinary least squares (OLS) regressions with robust standard errors in parentheses. Dependent variable is log real-effort output (EFFORT) during the experiment (in levels measured in grams). Treatment variables are sharecropping contract (C2), fixed-rent contract (C3) and ownership contract (C4), all binary indicators equal to one if the respective contract treatment was randomly assigned to a session and zero otherwise. The wage treatment (C1) is the reference contract. Significance level at 90(*), 95(**), 99(***) percent confidence. Controls A include household characteristics, controls B include production characteristics of cultivated plots, controls C include household shock indicators and controls D include indicators on government programs (see appendix B for a full description of variables). Differences between maximum subsample size n=460 and observations are due to missing values.

Table 16: Contract treatments and session mean real-effort output (white and yellow beans)

	(1) EFFORT	(2) EFFORT	(3) EFFORT	(4) EFFORT	(5) EFFORT	(6) EFFORT
C2	-0.0288	-0.0260	-0.0293	-0.0325	-0.0304	-0.0326
	{0.648}	{0.664}	{0.649}	{(0.608)}	{0.630}	{0.590}
	[0.639]	[0.665]	[(0.633)]	[0.597]	[0.621]	[0.581]
C3	0.0964	0.0973	0.0990	0.0980	0.0936	0.0936
	{0.135}	{0.114}	{0.134}	{0.184}	{0.198}	{0.208}
	[0.179]	[0.175]	[0.183]	[0.146]	[0.131]	[0.276]
C4	0.0715	0.0661	0.0781	0.0756	0.0696	0.0678
	{0.262}	{0.275}	{0.232}	{0.239}	{0.275}	{0.267}
	[0.306]	[0.277]	[0.276]	[0.287]	[0.318]	[0.276]
Controls A	no	yes	no	no	no	yes
Controls B	no	no	yes	no	no	yes
Controls C	no	no	no	yes	no	yes
Controls D	no	no	no	no	yes	yes
Observations	25	25	25	25	25	25

Notes: Ordinary least squares (OLS) regressions with p-values based on conventional standard errors in curly brackets and p-values based on robust standard errors in square brackets. Dependent variables are covariate-adjusted session effects (SESSION) from OLS regressions of real-effort output on a full set of session dummies and different sets of control variables A to E (not shown here). In (1) to (6) SESSION is regressed on the contract treatment of the respective session. Treatment variables are sharecropping contract (C2), fixed-rent contract (C3) and ownership contract (C4), all binary indicators equal to one if the respective contract treatment was randomly assigned to a session and zero otherwise. The wage treatment (C1) is the reference contract. Controls A include household characteristics, controls B include production characteristics of cultivated plots, controls C include household shock indicators and controls D include indicators on government programs (see appendix B for a full description of variables). Donald and Lang (2007) suggest using this two-stage procedure to avoid small sample bias in standard erros due to the possible presence of clustering effects when the treatment is randomized at the session level and the number of sessions is small. Given that all sessions are of equal size, no weights are attached to the second stage. Maximum session size is j=25.

Table 17: Contract treatments and session mean real-effort output (only white beans)

	(1) EFFORT	(2) EFFORT	(3) EFFORT	(4) EFFORT	(5) EFFORT	(6) EFFORT
C2	-0.0701	-0.0570	-0.0869	-0.0797	-0.0748	-0.0881
	{0.099}	{0.055}	{0.091}	{0.149}	{0.096}	{0.062}
	[0.062]	[0.027]	[0.021]	[0.084]	[0.050]	[0.013]
C3	0.0786	0.0920	0.0663	0.0817	0.0727	0.0785
	{0.069}	{0.006}	{0.182}	{0.131}	{0.104}	{0.090}
	[0.118]	[0.015]	[0.301]	[0.218]	[0.174]	[0.182]
C4	0.0601	0.0379	0.0634	0.0633	0.0537	0.0460
	{0.149}	{0.176}	{0.200}	{0.230}	{0.214}	{0.295}
	[0.122]	[0.154]	[0.128]	[0.177]	[0.191]	[0.239]
Controls A	no	yes	no	no	no	yes
Controls B	no	no	yes	no	no	yes
Controls C	no	no	no	yes	no	yes
Controls D	no	no	no	no	yes	yes
Observations	13	13	13	13	13	13

Notes: Ordinary least squares (OLS) regressions with p-values based on conventional standard errors in curly brackets and p-values based on robust standard errors in square brackets. Dependent variables are covariate-adjusted session effects (SESSION) from OLS regressions of real-effort output on a full set of session dummies and different sets of control variables A to E (not shown here). In (1) to (6) SESSION is regressed on the contract treatment of the respective session. Treatment variables are sharecropping contract (C2), fixed-rent contract (C3) and ownership contract (C4), all binary indicators equal to one if the respective contract treatment was randomly assigned to a session and zero otherwise. The wage treatment (C1) is the reference contract. Controls A include household characteristics, controls B include production characteristics of cultivated plots, controls C include household shock indicators and controls D include indicators on government programs (see appendix B for a full description of variables). Donald and Lang (2007) suggest using this two-stage procedure to avoid small sample bias in standard erros due to the possible presence of clustering effects when the treatment is randomized at the session level and the number of sessions is small. Given that all sessions are of equal size, no weights are attached to the second stage. Maximum session size for the subsample is j=13.

Table 18: Contract treatments and session mean real-effort output (only yellow beans)

	(1) EFFORT	(2) EFFORT	(3) EFFORT	(4) EFFORT	(5) EFFORT	(6) EFFORT
C2	0.0469	0.0404	0.0567	0.0504	0.0488	0.0514
	{0.317}	{0.367}	{0.227}	{0.319}	{0.322}	{0.308}
	[0.253]	[0.170]	[0.142]	[0.266]	[0.254]	[0.059]
C3	0.1486	0.1339	0.1592	0.1567	0.1494	0.1416
	{0.010}	{0.013}	{0.006}	{0.011}	{0.012}	{0.017}
	[0.024]	[0.037]	[0.014]	[0.030]	[0.030]	[0.045]
C4	0.1172	0.1321	0.1276	0.1195	0.1192	0.1392
	{0.029}	{0.014}	{0.019}	{0.037}	{0.032}	{0.020}
	[0.021]	[0.003]	[0.012]	[0.032]	[0.022]	[0.002]
Controls A	no	yes	no	no	no	yes
Controls B	no	no	yes	no	no	yes
Controls C	no	no	no	yes	no	yes
Controls D	no	no	no	no	yes	yes
Observations	12	12	12	12	12	12

Notes: Ordinary least squares (OLS) regressions with p-values based on conventional standard errors in curly brackets and p-values based on robust standard errors in square brackets. Dependent variables are covariate-adjusted session effects (SESSION) from OLS regressions of real-effort output on a full set of session dummies and different sets of control variables A to E (not shown here). In (1) to (6) SESSION is regressed on the contract treatment of the respective session. Treatment variables are sharecropping contract (C2), fixed-rent contract (C3) and ownership contract (C4), all binary indicators equal to one if the respective contract treatment was randomly assigned to a session and zero otherwise. The wage treatment (C1) is the reference contract. Controls A include household characteristics, controls B include production characteristics of cultivated plots, controls C include household shock indicators and controls D include indicators on government programs (see appendix B for a full description of variables). Donald and Lang (2007) suggest using this two-stage procedure to avoid small sample bias in standard erros due to the possible presence of clustering effects when the treatment is randomized at the session level and the number of sessions is small. Given that all sessions are of equal size, no weights are attached to the second stage. Maximum session size for the subsample is j=12.

Table 19: Contract treatments and session mean real-effort output excluding session 7 and 13 (white and yellow beans)

	(1) EFFORT	(2) EFFORT	(3) EFFORT	(4) EFFORT	(5) EFFORT	(6) EFFORT
C2	-0.0048	-0.0019	-0.0040	–0.0090	-0.0066	-0.0105
	{0.945}	{0.978}	{0.955}	{0.898}	{0.925}	{0.876}
	[0.944]	[0.978]	[0.953]	[0.896]	[0.924]	[0.875]
C3	0.1140	0.1114	0.1180	0.1167	0.1120	0.1081
	{0.099}	{0.092}	{0.095}	{0.093}	{0.105}	{0.105}
	[0.139]	[0.150]	[0.137]	[0.141]	[0.151]	[0.176]
C4	0.0891	0.0801	0.0971	0.0943	0.0880	0.0823
	{0.192}	{0.217}	{0.164}	{0.170}	{0.196}	{0.210}
	[0.235]	[0.230]	[0.206]	[0.213]	[0.240]	[0.230]
Controls A	no	yes	no	no	no	yes
Controls B	no	no	yes	no	no	yes
Controls C	no	no	no	yes	no	yes
Controls D	no	no	no	no	yes	yes
Observations	23	23	23	23	23	23

Notes: Ordinary least squares (OLS) regressions with p-values based on conventional standard errors in curly brackets and p-values based on robust standard errors in square brackets. Dependent variables are covariate-adjusted session effects (SESSION) from OLS regressions of real-effort output on a full set of session dummies and different sets of control variables A to E (not shown here). In (1) to (6) SESSION is regressed on the contract treatment of the respective session. Treatment variables are sharecropping contract (C2), fixed-rent contract (C3) and ownership contract (C4), all binary indicators equal to one if the respective contract treatment was randomly assigned to a session and zero otherwise. The wage treatment (C1) is the reference contract. Controls A include household characteristics, controls B include production characteristics of cultivated plots, controls C include household shock indicators and controls D include indicators on government programs (see appendix B for a full description of variables). Donald and Lang (2007) suggest using this two-stage procedure to avoid small sample bias in standard erros due to the possible presence of clustering effects when the treatment is randomized at the session level and the number of sessions is small. Given that all sessions are of equal size, no weights are attached to the second stage. Maximum session size is j=25.

Table 20: Order of experiments and session mean real-effort output (white and yellow beans, unweighted regressions)

	(1) EFFORT	(2) EFFORT	(3) EFFORT	(4) EFFORT	(5) EFFORT	(6) EFFORT
ORDER	-0.0043	0.0000	-0.0000	-0.0036	-0.0057	0.0021
	(0.0177)	(0.0166)	(0.0179)	(0.0178)	(0.0178)	(0.0167)
Controls A	no	yes	no	no	no	yes
Controls B	no	no	yes	no	no	yes
Controls C	no	no	no	yes	no	yes
Controls D	no	no	no	no	yes	yes
Mean (s.d.) dependend variable	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)
Observations	500	497	499	500	500	497
R-squared	0.0001	0.1649	0.0123	0.0143	0.0111	0.1927

Notes: Ordinary least squares (OLS) regressions with robust standard errors in parentheses. Dependent variable is log real-effort output (EFFORT) during the experiment (in levels measured in grams). Order of experiment (ORDER) is equal to one if the risk attitude experiment (used in a different paper) was conducted first in a session and equal to zero if the real-effort experiment was conducted second in a session. To avoid systematic spillovers from one experiment into the other, the order of the experiment was randomized across sessions. Under the null hypothesis of no spillover order of experiment does not have a significant effect on EFFORT. Significance level at 90(*), 95(**), 99(***) percent confidence. Controls A include household characteristics, controls B include production characteristics of cultivated plots, controls C include household shock indicators and controls D include indicators on government programs (see appendix B for a full description of variables). Differences between maximum sample size n=500 and observations are due to missing values.

Table 21: Order of experiments and real-effort output (white and yellow beans, weighted regressions)

	(1) EFFORT	(2) EFFORT	(3) EFFORT	(4) EFFORT	(5) EFFORT	(6) EFFORT
ORDER	-0.0163	-0.0134	-0.0134	-0.0153	-0.0169	-0.0139
	(0.0189)	(0.0180)	(0.0192)	(0.0188)	(0.0190)	(0.0182)
Controls A	no	yes	no	no	no	yes
Controls B	no	no	yes	no	no	yes
Controls C	no	no	no	yes	no	yes
Controls D	no	no	no	no	yes	yes
Mean (s.d.) dependend variable	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)
Observations	500	497	499	500	500	497
R-squared	0.0017	0.1528	0.0095	0.0158	0.0106	0.1786

Notes: Ordinary least squares (OLS) regressions with robust standard errors in parentheses. Dependent variable is log real-effort output (EFFORT) during the experiment (in levels measured in grams). Order of experiment (ORDER) is equal to one if the risk attitude experiment (used in a different paper) was conducted first in a session and equal to zero if the real-effort experiment was conducted second in a session. To avoid systematic spillovers from one experiment into the other, the order of the experiment was randomized across sessions. Under the null hypothesis of no spillover order of experiment does not have a significant effect on EFFORT. Significance level at 90(*), 95(**), 99(***) percent confidence. Controls A include household characteristics, controls B include production characteristics of cultivated plots, controls C include household shock indicators and controls D include indicators on government programs (see appendix B for a full description of variables). Differences between maximum sample size n=500 and observations are due to missing values.

Table 22: Contract treatments and real-effort output for sharecroppers in treatment 2 (white and yellow beans, unweighted regressions)

	(1) EFFORT	(2) EFFORT	(3) EFFORT	(4) EFFORT	(5) EFFORT	(6) EFFORT
C2	0.0004	0.0091	0.0014	0.0031	-0.0017	0.0011
	(0.0316)	(0.0280)	(0.0340)	(0.0315)	(0.0316)	(0.0294)
C3	0.0967***	0.0960***	0.0940***	0.0967***	0.0941***	0.0904***
	(0.0249)	(0.0237)	(0.0264)	(0.0246)	(0.0250)	(0.0251)
C4	0.0726***	0.0632***	0.0702***	0.0715***	0.0698***	0.0601**
	(0.0253)	(0.0230)	(0.0260)	(0.0250)	(0.0253)	(0.0235)
SHARECROP	0.0035	0.0050	-0.0078	0.0092	0.0054	-0.0018
	(0.0200)	(0.0194)	(0.0254)	(0.0200)	(0.0203)	(0.0248)
C2*SHARECROP	-0.0659*	-0.0738**	-0.0703*	-0.0736*	-0.0612	-0.0669*
	(0.0387)	(0.0356)	(0.0405)	(0.0388)	(0.0389)	(0.0362)
Controls A	no	yes	no	no	no	yes
Controls B	no	no	yes	no	no	yes
Controls C	no	no	no	yes	no	yes
Controls D	no	no	no	no	yes	yes
Mean (s.d.) dependend variable	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)
Observations	499	497	499	499	499	497
R-squared	0.0717	0.2283	0.0773	0.0846	0.0769	0.2457

Notes: Ordinary least squares (OLS) regressions with robust standard errors in parentheses. Dependent variable is log real-effort output (EFFORT) during the experiment (in levels measured in grams). Treatment variables are sharecropping contract (C2), fixed-rent contract (C3) and ownership contract (C4), all binary indicators equal to one if the respective contract treatment was randomly assigned to a session and zero otherwise. The wage treatment (C1) is the reference contract. SHARECROP is binary indicator equal to one if participants sharecrop in real life and zero if they do not and C2*SHARECROP is an interaction term. Significance level at 90(*), 95(**), 99(***) percent confidence. Controls A include household characteristics, controls B include production characteristics of cultivated plots, controls C include household shock indicators and controls D include indicators on government programs (see appendix B for a full description of variables).

Table 23: Contract treatments and real-effort output for sharecroppers in treatment 3 (white and yellow beans, unweighted regressions)

	(1) EFFORT	(2) EFFORT	(3) EFFORT	(4) EFFORT	(5) EFFORT	(6) EFFORT
C2	-0.0294	-0.0247	-0.0306	-0.0309	-0.0295	-0.0302
	(0.0255)	(0.0237)	(0.0269)	(0.0256)	(0.0253)	(0.0247)
C3	0.0651**	0.0617**	0.0652**	0.0616**	0.0633**	0.0562*
	(0.0307)	(0.0283)	(0.0315)	(0.0306)	(0.0307)	(0.0292)
C4	0.0742***	0.0648***	0.0730***	0.0730***	0.0712***	0.0619***
	(0.0252)	(0.0230)	(0.0260)	(0.0250)	(0.0253)	(0.0236)
SHARECROP	-0.0296	-0.0314	-0.0376	-0.0280	-0.0258	-0.0331
	(0.0203)	(0.0191)	(0.0254)	(0.0200)	(0.0211)	(0.0246)
C3*SHARECROP	0.0735**	0.0804**	0.0689*	0.0818**	0.0715*	0.0818**
	(0.0369)	(0.0354)	(0.0374)	(0.0378)	(0.0370)	(0.0368)
Controls A	no	yes	no	no	no	yes
Controls B	no	no	yes	no	no	yes
Controls C	no	no	no	yes	no	yes
Controls D	no	no	no	no	yes	yes
Mean (s.d.) dependend variable	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)
Observations	499	497	499	499	499	497
R-squared	0.0728	0.2293	0.0772	0.0859	0.0784	0.2481

Notes: Ordinary least squares (OLS) regressions with robust standard errors in parentheses. Dependent variable is log real-effort output (EFFORT) during the experiment (in levels measured in grams). Treatment variables are sharecropping contract (C2), fixed-rent contract (C3) and ownership contract (C4), all binary indicators equal to one if the respective contract treatment was randomly assigned to a session and zero otherwise. The wage treatment (C1) is the reference contract. SHARECROP is binary indicator equal to one if participants sharecrop in real life and zero if they do not and C3*SHARECROP is an interaction term. Significance level at 90(*), 95(**), 99(***) percent confidence. Controls A include household characteristics, controls B include production characteristics of cultivated plots, controls C include household shock indicators and controls D include indicators on government programs (see appendix B for a full description of variables).

Table 24: Contract treatments and real-effort output for sharecroppers in treatment 2 (white and yellow beans, weighted regressions)

	(1) EFFORT	(2) EFFORT	(3) EFFORT	(4) EFFORT	(5) EFFORT	(6) EFFORT
C2	0.0037	0.0113	0.0066	0.0085	0.0002	0.0051
	(0.0338)	(0.0300)	(0.0365)	(0.0333)	(0.0337)	(0.0313)
C3	0.1307***	0.1279***	0.1307***	0.1317***	0.1284***	0.1234***
	(0.0264)	(0.0251)	(0.0279)	(0.0260)	(0.0265)	(0.0265)
C4	0.0457*	0.0438*	0.0447	0.0473*	0.0437	0.0407
	(0.0276)	(0.0257)	(0.0284)	(0.0267)	(0.0276)	(0.0263)
SHARECROP	0.0127	0.0106	0.0043	0.0199	0.0164	0.0071
	(0.0213)	(0.0207)	(0.0267)	(0.0213)	(0.0219)	(0.0263)
C2*SHARECROP	-0.0608	-0.0710**	-0.0679*	-0.0715*	-0.0563	-0.0694*
	(0.0392)	(0.0360)	(0.0411)	(0.0394)	(0.0389)	(0.0368)
Controls A	no	yes	no	no	no	yes
Controls B	no	no	yes	no	no	yes
Controls C	no	no	no	yes	no	yes
Controls D	no	no	no	no	yes	yes
Mean (s.d.) dependend variable	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)
Observations	499	497	499	499	499	497
R-squared	0.0903	0.2370	0.0956	0.1048	0.0955	0.2535

Notes: Ordinary least squares (OLS) regressions with robust standard errors in parentheses. Dependent variable is log real-effort output (EFFORT) during the experiment (in levels measured in grams). Treatment variables are sharecropping contract (C2), fixed-rent contract (C3) and ownership contract (C4), all binary indicators equal to one if the respective contract treatment was randomly assigned to a session and zero otherwise. The wage treatment (C1) is the reference contract. SHARECROP is binary indicator equal to one if participants sharecrop in real life and zero if they do not and C2*SHARECROP is an interaction term. Significance level at 90(*), 95(**), 99(***) percent confidence. Controls A include household characteristics, controls B include production characteristics of cultivated plots, controls C include household shock indicators and controls D include indicators on government programs (see appendix B for a full description of variables).

Table 25: Contract treatments and real-effort output for sharecroppers in treatment 3 (white and yellow beans, weighted regressions)

	(1) EFFORT	(2) EFFORT	(3) EFFORT	(4) EFFORT	(5) EFFORT	(6) EFFORT
C2	-0.0240	-0.0215	-0.0238	-0.0247	-0.0256	-0.0267
	(0.0269)	(0.0252)	(0.0285)	(0.0267)	(0.0268)	(0.0261)
C3	0.0954***	0.0862***	0.0964***	0.0924***	0.0940***	0.0813***
	(0.0324)	(0.0301)	(0.0333)	(0.0324)	(0.0326)	(0.0313)
C4	0.0487*	0.0471*	0.0491*	0.0503*	0.0464*	0.0442*
	(0.0277)	(0.0257)	(0.0285)	(0.0268)	(0.0277)	(0.0264)
SHARECROP	-0.0222	-0.0299	-0.0281	-0.0203	-0.0168	-0.0292
	(0.0214)	(0.0205)	(0.0268)	(0.0213)	(0.0226)	(0.0262)
C3*SHARECROP	0.0823**	0.0973***	0.0834**	0.0911**	0.0798**	0.1012***
	(0.0366)	(0.0343)	(0.0368)	(0.0379)	(0.0368)	(0.0359)
Controls A	no	yes	no	no	no	yes
Controls B	no	no	yes	no	no	yes
Controls C	no	no	no	yes	no	yes
Controls D	no	no	no	no	yes	yes
Mean (s.d.) dependend variable	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)	4.93 (0.1977)
Observations	499	497	499	499	499	497
R-squared	0.0932	0.2410	0.0977	0.1076	0.0986	0.2587

Notes: Ordinary least squares (OLS) regressions with robust standard errors in parentheses. Dependent variable is log real-effort output (EFFORT) during the experiment (in levels measured in grams). Treatment variables are sharecropping contract (C2), fixed-rent contract (C3) and ownership contract (C4), all binary indicators equal to one if the respective contract treatment was randomly assigned to a session and zero otherwise. The wage treatment (C1) is the reference contract. SHARECROP is binary indicator equal to one if participants sharecrop in real life and zero if they do not and C3*SHARECROP is an interaction term. Significance level at 90(*), 95(**), 99(***) percent confidence. Controls A include household characteristics, controls B include production characteristics of cultivated plots, controls C include household shock indicators and controls D include indicators on government programs (see appendix B for a full description of variables).

B Variable definitions

This study employs both experimental and observational data. The variables used in the empirical analysis are described here.

Outcome variable. The outcome variable of interest is exercised EFFORT in a real-effort experiment. The variables is measured in grams and comes from weighting participants sorted beans on a scale after the real-effort experiment.

Treatment variables. The treatment variables of interest are different contracts commonly found in Ethiopian agriculture. These are binary indicators equal to one if a contract was a wage treatment (C1) and zero otherwise, equal to one if contract was a sharecropping treatment (C2) and zero otherwise, equal to one if contract was a fixed-rent treatment (C3) and zero otherwise and equal to one if a contract was an ownership treatment (C4). Contracts were randomly assigned to the different sessions.

Controls A. Controls A include a number of characteristic on the participant's household such as a binary indicator equal to one if the head is female and zero otherwise (GENDER), the age of the head (AGE), the age of the head squared (AGESQ), whether the household is member in a socio-political organization such as a cooperative (MEMBER), a binary indicator equal to one if head is literate and zero otherwise (LITERATE) and a count of household members (HHSIZE).

Controls B Controls B includes a number of general production characteristics as well as as land arrangement-specific characteristics such as whether households had a sharecropping arrangement (SHARECROPPING), fixed-rent arrangement (RENT) or employ wage labor (WAGE) on any of the plots they are cultivating during the last cropping season.

Further, we include whether households applied fertilizer (FERTILIZER), used a plough (PLOUGH), a saddle (SADDLE) or a threshing machine (THRESHING) on own plots or plots they sharecropped or rented-in. These variables differ for plots sharecropped and rented-in.

Controls C. Controls C includes variables on whether households were affected by a drought (DROUGHT), crop pest (PEST), livestock death (LDEATH), a substantial drop in sales prices for their agricultural produce (DROPSALES), the substantial increase in the price of food households have to buy (PRICEFOOD), whether a household member was affected by a health shock (HEALTH) and whether households were affected by sharp increases in the price for inputs (PRICEINPUT). All of which are binary indicators.

Controls D. Controls D include variables on whether households benefited from government programs such as the sustainable land management program (SLM), the infrastructure program (INFRASTRUCTURE), the savings and credit program (SAVINGS) and whether households belong to a self-help group (SELFHELP). All of which are binary indicators.

Other. I also include an expenditure aggregate. It includes expenditures on food, energy and transportation, education, health, social events, livestock, housing and other non-food expenditures. I calculate household per-capita expenditure (PCEXPENDITURE) to examine correlations between real contracts

and welfare. Following Deaton (1997) household members differing in age also receive different weights in the calculation of the per-capita aggregate: Children between 0 and 5 years of age receive a weight of 0.4 adults, children between 6 and 11 receive the weight of 0.5 adults. All household members are considered adults and receive a weight of 1. This variable is only used in table 1. Further, order of experiments (ORDER) is a binary indicator equal to one if the risk attitude game was conducted first and zero if the real-effort experiment was conducted second. Finally, SESSION denotes session effects estimated at the first-stage of the procedure suggested by Donald and Lang (2007).

C Web appendix

This study is accompanied by a comprehensive web appendix which allows for the reproduction of all tables presented here. It can be downloaded from:

http://froelich.vwl.uni-mannheim.de/2690.0.html. (link not active yet)

It contains a Stata data file with all variables used in the analysis (**master.dta**), a Stata Do-File which reproduces the empirical results presented in this paper (**generate_regressions.do**) and a detailed description of all variables employed in the analysis (see also Appendix C on variables). It further contains the survey instrument (**hhq.pdf**) which was used for collecting data on land-use, output by cropping season, capital inputs, labor inputs (all at the plot level) as well as household characteristics, participation in government programs, shocks, credit taking, expenditures and assets (all at the household level). It also contains a semi-structured interview on land arrangements (that is, contracts). It further contains the schedule for data collection sessions (**schedule.pdf**) and the experimental protocol (**real_effort.pdf**).

D Major revision to paper

From the Dec 2011 to Jan 2012 version the paper underwent the following major revisions:

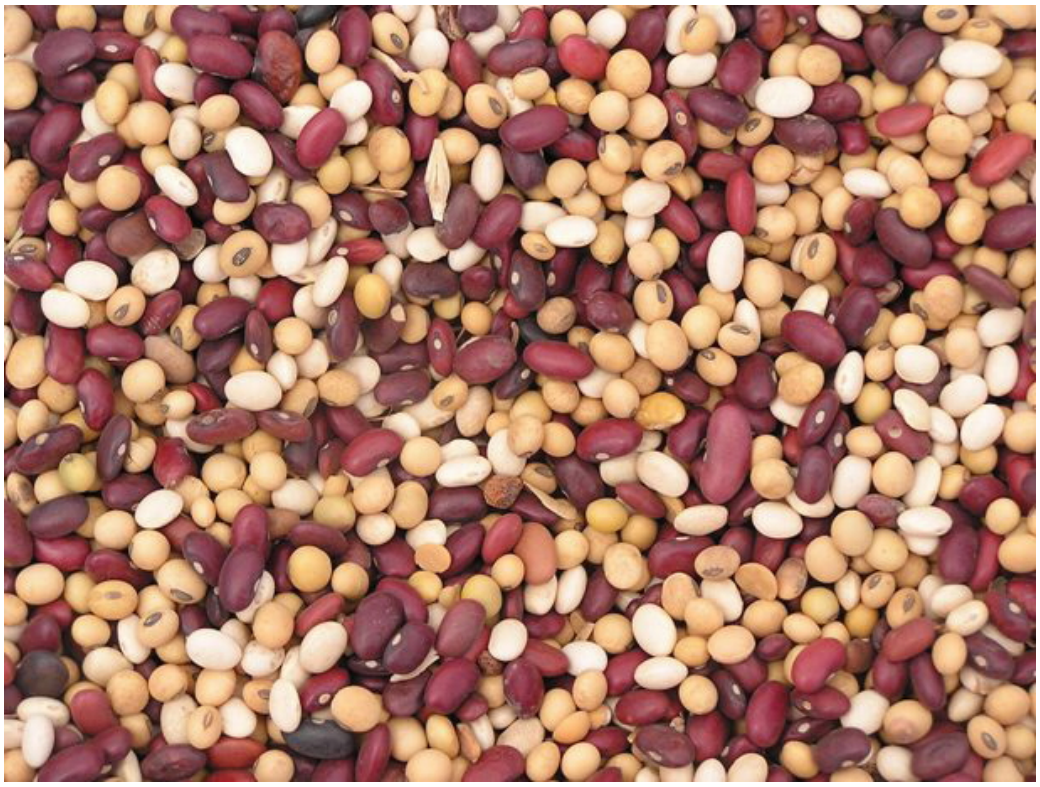
- Introduction and abstract were completely revised.
- The outcome variable of interest, real-effort output, is now measured in logs in the regression analysis.
- I rerun regressions with population weights (see table 14 to 19).
- I check for clustering effects. Following the two-stage procedure by Donald and Lang (2007) I try to avoid small sample bias in standard errors due to the possible presence of clustering effects when the treatment is randomized at the session level and the number of sessions is small (see table 20 to 25).

- I check whether the randomization of the order of the two experiments conducted in terms of the data collection helped avoid systematic spillover from one experiment into the other (see table 26 to 31).

From the Jan 2012 to the March 2012 version the paper underwent the following major revisions:

- A recheck of the do-file showed that not all intended control variables were included in the regressions. Including all of the intended control variables leads to marginal changes in the estimated treatment effects for regression including controls.
- A typo from a weight from one of the sessions was removed which slightly changes the point estimates for the sharecropping treatment in weighted regressions.

E Pictures from Experiment



Red, yellow and white beans for the real-effort task. For the real-effort task the enumarators asked participants to pick as many beans of a particular color as they can during a 15 minutes time period. I randomized the picking of yellow and white beans across sessions, but not the picking of red beans. Yellow and white beans were virtually alike in size. White beans turned out to weight a little more. However, red beans were bigger and heavier than yellow and white beans and therefore not used in the real-effort task (see table 2 for descriptive statistics on outcomes by yellow and white beans). Given the slight difference in weight between yellow and white beans I also show that the crucial empirical findings from the total sample also hold in the respective subsample (see tables 9 and 10).

Experimental equipment. Each participant received a flat, blue bucket filled with 2.1 kilos of beans in three different colors (700 grams of each) and an empty red bucket (without cover - unlike on the picture). Participants picked beans of one particular color from the blue bucket and dropped it into the red bucket. After the end of the real-effort task the beans from the red bucket were filled into the transparent small bowl and weighted. Payoffs were made according to the achieved weight of collected beans (the real-effort output). The Payoff depended on both the payoff scheme and achieved weight in grams (except for the wage contract were the payoff was fixed for all levels of real-effort output).

Seating order during experiment. During the real-effort task participants were sitting on the ground. The enumerators stood behind them on the inside of the circle. To allow for a big circle the experiment took place on the outside. To make sure that participants do not see each other exercising the task they turned around with their backs facing the inside of the circle. To make sure that participants do not hear each other exercising the task the bottom of the bucket was filled with a paper towel. In addition, we turned on music. During the task enumerators stood behind participants to make sure they do not turn their heads. Time was stopped with stopwatch.