Contract Farming as Partial Insurance*

Marc F. Bellemare† Yu Na Lee‡ Lindsey Novak§

September 25, 2020

*We thank seminar audiences at the University of Florida, the University of Georgia, the University of Göttingen, the University of Guelph, Kyoto University, as well as conference participants at the September 2016 FAO Rural Transformation, Agricultural, and Food System Transition conference, the 2017 Centre for the Study of African Economies conference, the 2017 Canadian Agricultural Economics Society annual meetings, and the 2017 Agricultural and Applied Economics Association annual meetings. We also thank two anonymous reviewers for useful comments and suggestions as well as EPAR at the University of Washington for making their Tanzania LSMS code publicly available. Bellemare is grateful to NIFA for funding this work through grant MIN-14-061—"Smallholder Participation in Agricultural Value Chains: Evidence from Madagascar." All remaining error are ours.

†Corresponding Author and Northrop Professor, Department of Applied Economics, University of Minnesota, 1994 Buford Ave, Saint Paul, MN 55108, United States, mbellema@umn.edu.

‡Assistant Professor, Department of Food, Agricultural & Resource Economics, University of Guelph, 50 Stone Road East, Guelph, ON N1G 2W1, Canada, ylee13@uoguelph.ca.

§Assistant Professor, Department of Economics, Colby College, 5237 Mayflower Hill, Waterville, ME 04901, United States, lindsey.novak@colby.edu.

CONTRACT FARMING AS PARTIAL INSURANCE

Abstract

A core result of contract theory is that contracts can help transfer risk from one party to another, the latter insuring the former. We test this prediction and explore the mechanism behind it in the context of contract farming, the economic institution wherein a processor contracts the production of a commodity to a grower. Specifically, we look at whether participation in contract farming is associated with lower levels of income variability in a sample of 1,200 households in Madagascar. Relying on a framed field experiment aimed at eliciting respondent marginal utility of participation in contract farming for identification in a selection-on-observables design, we find that participation in contract farming is associated with a 0.20-standard deviation decrease in income variability. Using mediation analysis to look at the mechanism behind this finding, we find support for the hypothesis that fixed-price contracts—which transfer all price risk from the grower to the processor—explain the reduction in income variability associated with contract farming. Because the assumption that makes our selection-on-observables design possible also satisfies the conditional independence assumption, we estimate propensity score matching and doubly robust weighted regression estimators, the results of which show that our core results are robust and that participation in contract farming would likely be more beneficial for those households that do not participate than for those who do. Our findings thus support the notion that, in a context where formal insurance markets fail, contracts can serve as partial insurance mechanisms.

Keywords: Risk and Uncertainty, Insurance, Contracts, Contract Farming, Agricultural Value Chains

JEL Classification Codes: L24, O13, O14, Q12

1 Introduction

Multiple market failures often lie at the root of economic underdevelopment and persistent poverty. In most developing countries, one market failure of particular importance is that of the insurance market. Usually in those cases, the insurance market fails because information problems—adverse selection and moral hazard—are important enough that it is often simply not profitable to offer insurance against risks which are commonly thought of as insurable in developed countries.

Insurance market failures constrain welfare in two ways. First, they constrain current welfare in that they force individuals and households to sink valuable resources into self-insuring against those risks, however partially.[1] Second, they constrain future welfare in that they prevent those same individuals and households from making today the requisite investments—financial, in agricultural technology, in education, and so on—that might otherwise allow them to attain higher levels of welfare tomorrow.

Contract farming, the vertical coordination mechanism wherein a processor contracts out the production of an agricultural commodity to a grower (Bijman, 2008), can in theory serve as a partial insurance mechanism for rural households in developing countries. Following Grosh (1994), contract farming can help resolve insurance market failures by insuring growers against

[1]Insurance can be full or partial. In the former case, the entirety of a risk is insured, and the insured party receives full compensation for its loss in case of an adverse event. In the latter case, only a fraction of a risk is insured, and the insured party receives less-than-full compensation for its loss in case of an adverse event.

price risk in cases where the processor guarantees a fixed price as part of the contract. This can lead to more stable incomes which, according to expected utility theory, means higher levels of welfare for risk-averse growers. More broadly, since Stiglitz's (1974) exploration of how sharecropping contracts can in theory partially insure tenants against output risk, economists have known that contracts can often help resolve market failures.

Does contract farming help empirically resolve insurance market failures? Specifically, does contract farming serve as a partial insurance mechanism for growers by reducing the income variability they face? We answer this question by using survey data on 1,200 households in rural Madagascar, half of which participate in contract farming as growers. To help disentangle the potentially causal relationship flowing from participation in contract farming to income variability from the correlation between the two, we rely on a framed field experiment to elicit each respondent's willingness to pay (WTP) to participate in a hypothetical contract farming agreement.[2] Because that WTP proxies for a respondent's marginal utility of participating in contract farming, and thus captures variations in that marginal utility due to typically unobservable factors (e.g., ambiguity and risk preferences, expected returns, time preferences, entrepreneurial ability, managerial ability, technical ability, aspirations, and so on), we argue that controlling for a respondent's WTP to participate in contract farming lessens statistical endogeneity issues stemming from grower self-selection into contract farming. As in Bellemare

[2] See List (2011) on framed versus artefactual field experiments.

and Novak (2017), we use the WTP data in a selection-on-observables design. To assess the robustness of our regression results, and given that our selection-on-observables design relies on an assumption that is similar to the conditional independence assumption made when estimating propensity score matching models, we also estimate propensity score matching models as well as doubly robust weighted regression estimators.

We find that participation in contract farming is associated with a decrease of about 0.20-standard deviations in the average household's income variability. Looking into the potential mechanisms underlying this finding, we use mediation analysis and a technique recently developed by Acharya et al. (2016) to look at whether our core finding is due to the presence of fixed-price contracts wherein the processor offers the grower a guaranteed fixed price. Our results suggest that (i) contract farming serves as a partial insurance mechanism for those households that choose to participate as growers,[3] and that (ii) this happens because fixed-price contracts transfer price risk from growers to processors.

There is a long, well-established empirical literature dating back to the early 1990s looking at the impacts of contract farming on the welfare of growers (Bellemare and Bloem, 2018). The bulk of that literature, however, looks at the effects of participation in contract farming on the *level* of income of

[3]We talk of partial insurance because even though our results support the hypothesis that contract farming helps growers insure against price risk, their income from contract farming depends on both the price they receive for the crops they grow under contract as well as on the quantity produced, and they still face production risk.

participating households (see, for instance, Glover, 1990; Singh, 2002; Warning and Key, 2002; Kumar and Kumar, 2008; Sharma, 2008; Maertens and Swinnen, 2009; Miyata et al., 2009; Jones and Gibbon, 2011; Bellemare, 2012; Mwambi et al., 2016; Wainaina et al., 2014; Wang et al., 2014; and Briones, 2015) or some variant thereof (Raynolds, 2002; Simmons et al., 2005; Begum, 2006; Minten et al., 2009; Bolwig et al., 2009; Narayanan, 2014; and Trifković, 2016). Beyond proximate outcomes like income and closely related variables (e.g., farm profits or farm revenue), however, the effect of participation in contract farming has only been documented for a handful of more distal outcomes such as the demand for women's labor (Raynolds, 2002), employment opportunities for women (Singh, 2002), happiness (Dedehouanou et al., 2013), food security (Bellemare and Novak, 2017; Montalbano et al., 2018), or technical efficiency (Mishra et al., 2018a). Minten et al. (2009) look at income variability, but their data lack a proper comparison group, and so they rely instead on an external data source for comparison. Michelson et al. (2012), for their part, find that relative to growers contracting with a domestic retail chain, Walmart growers in Nicaragua experience lower levels of price volatility, but their results focus on prices rather than on income variability. Additionally, Michelson et al. compare growers with those non-growers who have left the supply chain, which might skew the comparison relative to a control group selected randomly from the full population of non-growers, former growers, and never growers.

Our contribution is thus twofold. First, we contribute to the agricultural and development economics literatures by providing evidence that participation in agricultural value chains (Du et al., 2016; Lu et al., 2016; and Zilberman et al., 2017; Lu and Reardon, 2018) via the institution of contract farming can serve as a partial insurance mechanism for rural households in developing countries, and we do so with a considerable amount of external validity given that our data cover six regions of Madagascar, over a dozen different crops, and a number of different processors.[4] Second, and more importantly, we contribute to the literature on applied contract theory by documenting that the likely mechanism whereby contract farming serves as a partial insurance mechanism is via contracts that transfer output price risk from the grower to the processor.[5]

The remainder of this paper is organized as follows. Section 2 lays out a

[4]Two previous studies (Bellemare 2012, and Bellemare and Novak 2017) used the data we use in this paper to study contract farming, so an anonymous reviewer has asked us to clarify how this paper contributes to the literature. Recall that the two previous studies were about income (Bellemare 2012) and food security (Bellemare and Novak 2017) as measured by the duration of the hungry season experienced before harvest—that is, with two proxies for welfare that have to do with a household's level of income (i.e., the first moment of the income distribution) either directly (in the case of income) or indirectly (in the case of food security, which captures how much income is saved from one harvest to the weeks leading up to the next harvest). By contrast, this paper is focused on income variability (i.e., the second moment of the income distribution). Though our theoretical framework in particular (and contract theory in general) is clear that contracts can and should be used as risk-sharing instruments, there is no reason why this should be true in practice. Likewise, though the efficient markets hypothesis predicts that assets with higher returns should also have higher variance, we find that this is not true in the case of the contracts we consider here, if one considers them as financial instruments. Finally, unlike this study, neither Bellemare (2012) nor Bellemare and Novak (2017) study the mechanisms behind their findings.

[5]On the consequences of price risk for the welfare of producers, see the theoretical studies by Baron (1970) and Sandmo (1971), the observational studies by Barrett (1996) and Bellemare et al. (2013), and recent experimental work by Bellemare, Lee, and Just (2020).

simple theoretical framework showing the mechanisms whereby participation in contract farming can serve as a partial insurance mechanism for participating households. In section 3, we present the empirical framework we rely on to study the effects of participation in contract farming on income variability, paying particular attention to our identification strategy. Section 4 presents the data and discusses some descriptive statistics. In section 5, we present our empirical results. Section 6 concludes with policy implications and with some directions for future research.

2 Theoretical Framework

We consider the maximization problem of an individual grower of an agricultural commodity who is considering whether and how much to produce under contract for a processor. As such, we are not concerned with the processor's decision of whether or not to contract the production of the agricultural commodity to growers or to produce it in-house. Rather, we take as given the processor's decision to contract out the production of the agricultural commodity.

Assume that a representative producer growing a single crop has a von Neumann-Morgenstern utility function $U(\cdot)$ defined over profit π. The function $U(\pi)$ is twice continuously differentiable, strictly increasing, and strictly concave in profit, i.e., $U_\pi > 0$ and $U_{\pi\pi} < 0$.[6] Let p be a piece rate, i.e., the

[6]For any function $f(\cdot)$, we let f_k and f_{kk} denote the first and the second derivatives of $f(\cdot)$ with respect to k, respectively.

price at which the producer can sell each unit of his crop q at market after harvest; this piece rate is a random variable.

The producer can choose to participate in contract farming by agreeing to sell a fraction $\alpha \in [0, 1]$ of his crop to a processor who will pay the certain fixed price $\overline{p} > 0$ for each unit of q. In that case, the producer's profit is such that

$$\pi = \{(1 - \alpha)p + \alpha\overline{p}\}q - TC(q, c, \alpha), \tag{1}$$

where $TC(q, c, \alpha)$ denotes the total cost of producing output q when the input cost of producing each unit is c. The total cost varies according to α as well due to the costs associated with contract farming, such as transaction and compliance costs. The function $TC(q, c, \alpha)$ is twice continuously differentiable and strictly increasing in q and c, i.e., $TC_x > 0$ for $x = q, c$.

Here, we assume that the total cost function $TC(q, c, \alpha)$ is nonseparable in terms of q, c, and α for the sake of simplicity. The results presented in this section do not change when we express the total cost as the sum of fixed and variable costs, or when we express fixed, variable, or total costs as a linear function of q, c, and α.

Because the market price p is a random variable,[7] the producer's expected

[7]In reality, risk exists in quantity and input prices as well as output price due to various factors such as shocks in weather, yield, and demand, and other random shocks. Moreover, price is determined through an endogenous process in which shocks to supply and demand are filtered through to prices by general equilibrium. In this theoretical framework, however, we focus on output price as the random variable in order to keep the analysis simple and tractable.

profit is such that

$$E(\pi) = \int_0^\infty \left[\{(1-\alpha)p + \alpha\bar{p}\}q - TC(q,c,\alpha)\right] f(p)dp, \tag{2}$$

where $E(\cdot)$ denotes an expectation and $f(p)$ denotes a non-degenerate probability distribution function of p, with the expected value $E(p) = \mu$. Similarly, the variance of the producer's profit is such that

$$Var(\pi) = \int_0^\infty \left[\{(1-\alpha)p + \alpha\bar{p}\}q - TC(q,c,\alpha) - E(\pi)\right]^2 f(p)dp. \tag{3}$$

$E(\pi)$ can be rewritten as

$$E(\pi) = \{(1-\alpha)\mu + \alpha\bar{p}\}q - TC(q,c,\alpha), \tag{4}$$

which means that

$$Var(\pi) = \int_0^\infty \left[(1-\alpha)(p-\mu)q\right]^2 f(p)dp. \tag{5}$$

The foregoing leads to the following proposition, which is our core testable hypothesis.

Proposition 1 *Under the assumptions made so far, participation in contract farming decreases the variance of a participating producer's profit. Moreover, given participation in contract farming, the higher the contract coverage* α*, the lower the variance of the producer's profit.*

Proof. First, let us compare the cases of not participating (i.e., $\alpha = 0$) and participating in contract farming ($0 < \alpha \leq 1$).

If $\alpha = 0$,

$$E(\pi|\alpha = 0) = \mu q - TC(q, c, 0), \text{ and} \tag{6}$$

$$Var(\pi|\alpha = 0) = \int_0^\infty \big[(p - \mu)q\big]^2 f(p)dp. \tag{7}$$

If $0 < \alpha \leq 1$,

$$E(\pi|0 < \alpha \leq 1) = \{(1 - \alpha)\mu + \alpha\bar{p}\}q - TC(q, c, \alpha), \text{ and} \tag{8}$$

$$Var(\pi|0 < \alpha \leq 1) = (1 - \alpha)^2 \int_0^\infty \big[(p - \mu)q\big]^2 f(p)dp. \tag{9}$$

Therefore, $Var(\pi|\alpha = 0) > Var(\pi|0 < \alpha \leq 1)$.

Next, given participation in contract farming, the change in $Var(\pi)$ according to α is such that

$$\frac{\partial Var(\pi)}{\partial \alpha} = -\int_0^\infty 2(1 - \alpha)\big[(p - \mu)q\big]^2 f(p)dp \leq 0. \tag{10}$$

The last inequality holds due to $0 < \alpha \leq 1$ given participation in contract farming. ■

The producer's maximization problem can be expressed as follows:

$$\max_{\alpha,q} EU(\pi) = \max_{\alpha,q} \int_0^\infty U\Big(\{(1 - \alpha)p + \alpha\bar{p}\}q - TC(q, c, \alpha)\Big) f(p)dp, \tag{11}$$

which leads to the following lemma.

Lemma 2 *If contract farming guarantees a price equal to the expected market price and the marginal cost of participating in contract farming is zero, a risk-averse producer will benefit from full coverage for a given level of production. That is, if $\bar{p} = \mu$ and $TC_\alpha(q, c, \alpha) = 0$ at any $0 \leq \alpha \leq 1$, the producer's choice of α will be equal to* 1.

Proof. Consider two choices: (i) full participation in contract farming ($\alpha = 1$) and (ii) no participation ($\alpha = 0$) or a partial participation ($0 < \alpha < 1$) given a level of production q. The contract guarantees $\bar{p} = \mu$. The producer will benefit from fully participating in contract farming if and only if:

$$EU[\pi|\alpha = 1] - EU[\pi|0 \leq \alpha < 1] > 0. \tag{12}$$

Note that the left-hand-side of expression 12 is equal to

$$= EU[\mu q - TC(q, c, 1)] - EU[\{(1 - \alpha)p + \alpha\bar{p}\}q - TC(q, c, \alpha)] \tag{13}$$

$$= EU[\{(1 - \alpha)\mu + \alpha\bar{p}\}q - TC(q, c, \alpha)] - EU[\pi] \tag{14}$$

$$= EU[E(\pi)] - EU[\pi] \tag{15}$$

$$= U[E(\pi)] - EU[\pi] > 0. \tag{16}$$

Expression 14 follows from $\bar{p} = \mu$ and $TC_\alpha(q, c, \alpha) = 0$ at any $0 \leq \alpha \leq 1$. The last expression follows from assuming that $U_{\pi\pi} < 0$ and by Jensen's inequality. ■

Relaxing the assumption that the marginal cost of contract farming is zero, which would be the case if there are additional costs associated with adhering to the demands of the processor, leads to the following proposition.

Proposition 3 *If the marginal cost of contract farming is nonnegative (i.e., $TC_\alpha(q, c, \alpha) \geq 0$), then the optimal level of coverage α^* is determined such that the marginal benefit and marginal cost from participating in contract farming are equalized.*

Proof. By the F.O.C. with respect to α, it must be that

$$E\Big[\frac{\partial U(\pi)}{\partial \alpha}\Big] = E\Big[\frac{\partial U(\pi)}{\partial \pi} \cdot \frac{\partial \pi}{\partial \alpha} + \frac{\partial U(\pi)}{\partial q} \cdot \frac{\partial q}{\partial \alpha}\Big] = 0, \tag{17}$$

where $\frac{\partial EU(\pi)}{\partial q} = 0$ at the optimal level of production. Therefore, it must be that

$$E\Big[\frac{\partial U(\pi)}{\partial \pi} \cdot \frac{\partial \pi}{\partial \alpha}\Big] = \int_0^\infty U_\pi(\pi)\big[(\bar{p} - p)q - TC_\alpha(q, c, \alpha^*)\big] f(p)dp = 0, \tag{18}$$

which means that, at the optimum,

$$(\bar{p} - p)q = TC_\alpha(q, c, \alpha^*) \tag{19}$$

given that $U_\pi > 0$. ■

In other words, in order for a producer to be better off choosing to participate in contract farming, the contract must pay a fixed price that is high enough to cover any cost borne by the producer because of his participation in the contract. That is, the difference between the fixed price guaranteed by contract farming and the market price ($\bar{p} - p$) is the premium paid by the processor to secure the level of production q. Though we do not formally test Proposition 3, we derive it here in order to explain why producers who choose to participate in contract farming do not go "all in" by choosing to cultivate the entirety of their plots under contract farming.

We summarize this section with three testable hypotheses that we aim to test empirically.

Hypothesis 1 *Income variability is lower for households that participate in contract farming than households that do not.*

Hypothesis 1 is derived directly from Proposition 1. Because our data set does not include information on input cost or producer profit, however, we focus on household income as a measure of welfare.

Hypothesis 2 *The greater the proportion of a household's plots is under a fixed-price contract, the lower the variability of that household's income.*

Hypothesis 2 is also based on Proposition 1 and concerns the mechanism whereby contract farming affects income variability. We use the proportion of a household's plots under a fixed price contract as a proxy for α, the level of contract coverage under a fixed price contract.

Hypothesis 3 *Fixed price contracts are the only mechanism whereby participation in contract farming reduces income variability.*

Hypothesis 3 goes one step further than Hypothesis 2 and allows us to rule out other potential causal pathways from contract farming to income variability.

3 Empirical Framework

We now discuss the empirical framework we use to study the impact of participation in contract farming on income variability. We begin this section by discussing how we build our measure of income variability—that is, our outcome of interest—for the remainder of this paper. We then move on to our estimation and identification strategies.

3.1 Measurement of Income Variability

One difficulty in answering the research question we pose is that we rely on cross-sectional data. Ideally, one would have longitudinal data at one's disposal to measure the variability of a household's income over time. That way, one could obtain for each household a measure of central tendency (e.g., the within-household mean or median) of that household's income in order to then estimate how far that household's typically lies on average from that measure of central tendency. For example, one could use longitudinal data to simply compute the standard deviation or the variance of a household's income over time.

Unfortunately, we know of no publicly available longitudinal data set that would allow studying the relationship between participation in contract farming and income variability. And even if such a data set were available, there is no guarantee that it would allow incorporating household fixed effects in an attempt to identify the relationship between participation in contract farming and income variability, as there is usually little movement in and out of contract farming from year to year at the household level.

In this sense, we wish to state in no uncertain terms that the measure of income variability we use below is a proxy for income variability. Future research should focus on collecting and analyzing panel data that cover a long enough time period to exhibit sufficient within-household variation in contract farming participation.[8]

For completeness, Appendix A shows how, in two well-known longitudinal developing-country data sets (i.e., the Tanzanian Living Standards Measurement Surveys and the Ethiopia Rural Household Survey), cross-sectional income variability is a good proxy for longitudinal income variability, and so we use the available former here as a proxy for the unavailable latter in our own data. Indeed, among rural Tanzanian and Ethiopian households, our measure of income variation, while not perfectly correlated, is highly predictive of within-household longitudinal income variation. The correlation coefficient is approximately 0.640 among Tanzanian households and 0.375 among Ethiopian households. Further, we show that in both Tanzania and

[8] In an earlier version of this paper, we used two additional proxies for income variability, and our results were robust to using any of these three proxies. For the sake of brevity and in response to an anonymous reviewer asking us to shorten this paper, we focus here only on a single proxy for income variability.

Ethiopia, these correlation coefficients are largest among households with a married, male, migrant household head whose landholdings are less than the median amount in the relevant sample.

To look at whether there is a relationship between participation in contract farming and income variability, we conduct a test of the null hypothesis that the heteroskedasticity of household income (measured in levels, per capita, and per adult equivalent for robustness) is no different between those households that participate in contract farming and those that do not. Indeed, since Engle's (1982) seminal contribution, economists have recognized that there is empirical content to heteroskedasticity in that heteroskedasticity, because it measures the variance around a regression line, measures volatility of that regression's outcome variable holding explanatory variables constant.

To do so, we first estimate the equation

$$\ln(y_i) = \alpha_0 + \underline{\beta}_0 \underline{x}_i + \gamma_0 D_i + \epsilon_{0i}, \tag{20}$$

where y_i denotes household i's income y_i, $\underline{x}$ is a vector of household-specific control variables,[9] D is an indicator variable for whether the household participates in contract farming, and ϵ is an error term with mean zero. The variables included in $\underline{x}$ are respondent marital status, her gender, whether she migrated into the village from elsewhere, her age, her education level, her

[9]Throughout this paper, underlines denote vectors.

years of agricultural experience, and her membership in a farm organization as well as the number of days that farming is forbidden for her, the value of her household's working capital and assets, the size of her household's landholdings, the size of her household, and her household's dependency ratio.

Our heteroskedasticity (H) measure is such that, for each household i, we compute

$$H_i = \widehat{\epsilon}_i^2, \tag{21}$$

where $\widehat{\epsilon}_i$ denotes the residual for household i, whose square we use as our measure of income variability in two distinct approaches.

To state it as clearly as possible, the assumption we make here is that the squared absolute distance between a household's income realization and the predicted income based on the control variables in $\underline{x}$ is representative of what that distance would be on average for the same household over time, in longitudinal data. Put differently, we assume here that equation 20 is correctly specified, and that equation 21 measures what is known as *pure heteroskedasticity* (Goldberger 1991).

To test whether income variability is the same across the sub-samples of households that participate in contract farming and households that do not, we conduct two heteroskedasticity tests. We first conduct an unconditional heteroskedasticity test, i.e., a t-test of the null hypothesis that $\overline{CH} = \frac{1}{N}\sum_{i=1}^{N} H_i$ does not differ between the sub-samples of households that participate in contract farming and those that do not; this is a test whose goal is to establish whether income variability is the same for those households that participate in contract farming and those households that do not.

Second, we use H_i as our dependent variable in a regression of H_i on the variable of interest (i.e., participation in contract farming), and the control variables, such that

$$H_i = \alpha_1 + \underline{\beta}_1 \underline{x}_i + \gamma_1 D_i + \epsilon_{1i}. \tag{22}$$

This is a test of conditional heteroskedasticity, since it conditions on more than just the treatment variable.

3.2 Estimation Strategy

This section discusses the three approaches—regression, matching, and the combination of the two—we use in order to study the relationship between participation in contract farming and income variability. In what follows, we closely follow the notation in Bellemare and Novak (2017).

3.2.1 Regression

Starting with the regression approach, our core estimable equation is equation 22 above. Our coefficient of interest is γ. If D were randomly assigned, γ would provide an estimate of the average treatment effect (ATE) of participating in contract farming on the proxy for income variability on the left-hand side of equation 22. Participation in contract farming, however, is not randomly assigned, and so we estimate the following version of equation 22:

$$H_i = \alpha_2 + \underline{\beta}_2 \underline{x}_i + \gamma_2 D_i + \underline{\delta}_2 \underline{r}_i + \epsilon_{2i}, \tag{23}$$

where H, $\underline{x}$, D, and ϵ are defined as before, but where $\underline{r}_i$ is a vector of variables capturing whether a respondent would be willing to pay various amounts of money in order to participate in contract farming. Since we use this vector of proxy measures of willingness to pay (WTP) to account for selection in contract farming, and thus in an attempt to identify the impact of participation in contract farming on income variability, we defer our discussion of it to the next subsection. Until then, note that the research design we rely on in this paper is a selection-on-observables (SOO) design.

3.2.2 Propensity Score Matching

The assumption that allows us to use an SOO design also allows us to make the conditional independence assumption in this context (Imbens, 2015), which means that we can use propensity score matching (PSM) methods to answer the research question posed in this paper. In this context, the use of PSM methods has two distinct advantages. First, it allows assessing the robustness of our regression results. Second, it allows estimating average treatment effects on the treated (ATT) and on the untreated (ATU), two measures which can be useful to inform economic policy in this context but which are not estimable from the regression in equation 23.

To use PSM methods, we proceed in two steps. First, we estimate the following probit:

$$D_i = \kappa + \underline{\lambda x}_i + \underline{\theta r}_i + \xi_i, \tag{24}$$

where the variables denote the same things as before. The estimated coefficients in equation 24 are then used to obtain a prediction $\widehat{D}$ of the dependent variable—the propensity score, which measures the likelihood that each individual observation i is treated, i.e., the likelihood that a household i participates in contract farming estimated on the basis of the covariates on the RHS of equation 24.

Second, we match households that participate in contract farming with households that do not on the basis of their propensity scores. To do so, we match with replacement and use the three nearest neighbors with a caliper size of 0.01 standard deviation.[10] For each of our proxies for income variability, we report the ATE, the ATT, and the ATU.

[10]In preliminary work, we also considered two other specifications: (i) one nearest neighbor with a caliper size of 0.01 standard deviation, and (ii) three nearest neighbors with a caliper size of 0.001 standard deviations. All three specifications gave qualitatively similar results, and so we only report one for brevity.

3.2.3 Doubly Robust Weighted Regression Estimators

Given that we argue in section 3.2.1 that we have an SOO research design and in section 3.2.2 that this also allows us to use PSM methods, an anonymous reviewer has suggested that we combine the two into a doubly robust (DR) weighted regression estimator. The DR estimator is a two-step procedure which consists in (i) estimating propensity scores as in section 3.2.2, and (ii) using those propensity scores as weights in a regression as in section 3.2.1 (Ho et al., 2007). In the words of Morgan and Winship (2015), this procedure gives the researcher two chances to get the estimates right, provided the propensity score-estimating equation in equation 24 is correctly specified.

3.3 Identification Strategy

Recall that we rely on a selection-on-observables design to look at the relationship between participation in contract farming and income variability. In this section, we first explain the framed field experiment used to elicit respondent WTP to participate in contract farming. Then, we go through the usual sources of statistical endogeneity—unobserved heterogeneity, reverse causality, and measurement error—and explain how well our approach addresses each one.

3.3.1 Experimental Setup

A contingent valuation experiment was run in the field that asked each respondent whether he would be willing to pay a randomly selected amount of money (hereafter, the bid b) in order to participate in a hypothetical contract farming agreement that would increase his income by 10 percent.

Each respondent's bid b_i was selected from the set $\{\$12.5, \$25, \$37.5, \$50, \$62.5, \$75\}$ with equal probability (i.e., with the throw of a fair six-sided die).[11] To give the reader some perspective, the average annual household income in our data was equal to about \$970, and so the bid could range anywhere from about 2 to 8 percent of that average household's annual income.

For each respondent i, we have a binary-choice (i.e., yes or no) answer to the bid b_i-specific question posed in the framed field experiment, but not for the five other potential bids. One immediate problem, then, is that a respondent is not asked whether she is willing to participate in contract farming at all levels of the bid variable. Indeed, for each respondent, we know whether she would be willing to participate in the hypothetical contract farming agreement only for the bid that was randomly drawn for her. Eliciting a response for just one bid is common in the contingent-valuation literature in order to avoid respondents anchoring their response on the previous level.

Strictly speaking, each respondent is presented with bid vector $\underline{b}_i$, which is such that $\underline{b}_i = \{b_{1i}, b_{2i}, b_{3i}, b_{4i}, b_{5i}, b_{6i}\}$. For each respondent i, only one

[11] Dollar amounts are reported here for ease of exposition. During fieldwork, respondents were presented with equivalent amounts stated in the local currency.

of the elements of $\underline{b}_i$ is equal to one, with all the others equal to zero. For instance, a respondent who rolls a four on the die throw used to randomize the bid would have a vector of bids $\underline{b}_i = \{0, 0, 0, 1, 0, 0\}$. If that respondent says "Yes" to the $50 bid that a die throw of four maps into, then her vector of WTP proxies $\underline{r}_i$ will have a one in the fourth position.

What remains to be determined is the value of her vector $\underline{r}_i$ in the other five positions. In order to ascribe respondent-specific values to bids a respondent is not asked (i.e., to recover what a respondent's answer would be for the other, unasked bids) we proceed as follows. If a respondent said "Yes" to a given bid value, we code the answer to every bid of lower value as a "Yes." So for our respondent who said "Yes" to paying $50, we assume that she would also be willing to pay $12.50, $25, and $37.50. For higher bid values, since we do not know whether the respondent would have said yes to those, we conservatively assume that she would say "No" to those. So for our respondent who said "Yes" to paying $50, we assume that she would not be willing to pay $62.50 and $75. Finally, for those respondents who said "No" to the bid they were presented with, we assume that they would say no to all possible bid levels.

To be clear, the foregoing assigns a lower bound on their proxy measure of WTP to participate in a contract farming agreement that would increase their income by 10 percent. By assuming that a respondent who says "Yes" to a given bid level would also agree to all lower bids, we impose that her bid is the lower bound on her WTP. By assuming that a respondent who says "No" to a given bid level would say no to all other bid levels, we use the information we have and assume that that respondent's WTP is equal to zero. The end result is a vector $\underline{r}_i$ that is such that

1. $\underline{r}_i$ is composed only of ones if the respondent said "Yes" to a bid of \$75.

2. $\underline{r}_i$ is composed only of zeros if the respondent said "No" to her bid.

3. $\underline{r}_i$ is composed of ones and zeros if the respondent said "Yes" to a bid value strictly less than \$75.

The vector $\underline{r}_i$ thus proxies for a respondent's WTP to participate in contract farming, and we use this vector as a control in our SOO design.[12]

An anonymous reviewer raised the concern that our respondent's WTP responses might suffer from confirmation bias in the sense that those responses would reflect our respondents' choices rather than their preferences. Chen (2008) and Chen and Risen (2010), however, show that psychologists' view that preferences reflect choices (instead of the economists' view that choices reflect preferences) is based on flawed studies, and they conclude that in the vast majority of cases, choices reflect preferences.

[12] In preliminary work for this paper, we also imputed what a respondent's answer would be for the unasked bids as follows. If a respondent was presented with bid b_{ij}, where j denotes any of the six possible bid levels, we linearly regress each unasked bid $b_{i,-j}$ on $\underline{x}$ and predicted $\widehat{r}_{i,-j}$. The end result was a vector $\underline{r}_i = (r_{i,j}, \widehat{r}_{i,-j})$ which summarizes (i) whether the respondent reports being willing to pay the bid amount randomly selected for him by the throw of a die to participate in the hypothetical contract farming agreement, and (ii) the likelihood that he would be willing to pay the other bid amounts. Our results are robust to using these imputations.

It is also legitimate to worry about whether nonparticipants in contract farming have a good understanding of the costs and benefits of participation in contract farming. For many treatment variables of interest in the applied microeconomics literature, participants and nonparticipants indeed have a very different understanding of what is at stake. Here, we are confident that this is not an issue: The villages respondents were sampled from are small villages that remain agrarian economies, so agriculture is our respondents' primary occupation, and people in those villages not only discuss agriculture all the time between themselves, but everyone tends to know one another. In this sense, contract farming is someone that one learns about from others in addition to learning by doing (Foster and Rosenzweig 1995). We now turn to how this vector reduces the bias that would result from grower self-selection into contract farming.

3.3.2 Identification

The identifying assumption we make is that the vector $\underline{r}_i$, which proxies for a respondent's WTP to participate in contract farming at all bid levels, accounts for a respondent's marginal utility of participating in contract farming, which allows controlling for selection into contract farming by purging the error term in equation of much of its correlation with the treatment variable in equation 23. In this section, we explain the reasoning behind this claim, which allows both adopting the SOO design just laid out in the regression context as well as assuming that conditional independence holds in the matching context.

Any identification strategy has to be judged based on how it fares relative to the three usual sources of statistical endogeneity, viz. (i) unobserved heterogeneity, (ii) reverse causality, and (iii) measurement error. Much of our discussion follows that in Bellemare and Novak (2017).

Unobserved Heterogeneity. Many of our respondents' characteristics are unobserved. When unobservable characteristics are correlated with variables on the RHS of the equation of interest (here, equation 23), estimated coefficients are biased. Here, because the vector $\underline{r}_i$ captures a respondent's marginal utility of participating in contract farming, many of the typically unobservable characteristics whose correlation with D would bias our estimate of γ in equation 23 (e.g., risk and ambiguity preferences, entrepreneurial and technical ability, etc.) are accounted for by shifts in a respondent's marginal utility of participating in contract farming. For example, suppose two respondents are identical, except for their entrepreneurial ability. The respondent whose entrepreneurial ability is higher might prefer starting a business to participate in contract farming; this would be reflected in his having a higher marginal utility of participating in contract farming relative to the other respondent, and this difference would be captured in different values of the vector $\underline{r}_i$ for the two respondents. A similar reasoning applies to other unobservable sources of variation in marginal utility of participating in contract farming which could be correlated with variables on the RHS of equation 23, which considerably lessens the problem of unobserved heterogeneity.

Reverse Causality. One way in which this would arise in cases where the prospect of getting partial insurance via reduced income variability would induce respondents to participate in contract farming. In such scenarios, our estimate of γ would be biased because of reverse causality flowing from income variability to participation in contract farming. In one sense, this is another version of the unobserved heterogeneity story. Indeed, assume once again that two respondents are identical, save for their willingness to participate in contract farming because of how they differ in their expectation that contract farming will serve to partially insure them. These different expectations would affect their marginal utility of participating in contract farming differentially which, again, the vector $\underline{r}_i$ would account for.

Measurement Error. This would arise in cases where our variable of interest, i.e., D, the dummy variable which measures whether respondents participate in contract farming or not, were measured with error. This is not a concern here, for three reasons. First, there is no incentive to lie about this, as there is no social stigma attached to participating in contract farming, nor is there a benefit to responding one way or the other. Second, there are no recall problems for this question, because respondents are fully aware of whether they participate in contract farming or not. Finally, the sampling frame was established with village leaders, who made two lists for their community, one of all the households that participated in contract farming and one of all the households that did not, from which enumerators randomly selected respondents. This served as an additional check that respondents accurately reported their participation status. If there is any measurement error, it occurs at random, and it should be so minimal as to be unlikely to cause much attenuation bias.[13]

Our identification strategy thus helps accounting for a number of sources of statistical endogeneity, but it is not perfect. The ideal research design would involve randomly assigning treatment households to participation in contract farming and control households to nonparticipation, but so far the only such example of randomized assignment in the context of contract farming is by Arouna et al. (2019). As always with observational data, however, it is best to exercise caution, and so it is best to treat our estimates of γ as suggestive of causality rather than causal.

4 Data and Descriptive Statistics

Bellemare (2012) and Bellemare and Novak (2017) rely on the data we use in this paper, and Bellemare and Lim (2018) discuss the same data in detail, so we dedicate only a limited amount of space to discussing the data. The reader interested in knowing more about the details of data collection, descriptive statistics, the features of the contract farming agreements we study, and

[13] Another threat to identification would be violations of the stable unit treatment value assumption (SUTVA) in the form of spillovers from one household's contract farming participation status to another household's income variability. Though this is in theory possible, there is unfortunately nothing we can do about such SUTVA violations in this study.

so on is encouraged to read Bellemare and Lim (2018). Though we report descriptive statistics for the variables we use in our analysis, we will not expend any time discussing them, as the aforementioned articles do that.

The data were collected during the latter half of 2008 in 12 communes across six regions of Madagascar, with two communes sampled per region. The data cover 1,200 households, half of which participate in contract farming and half of which do not. Regions were selected on the basis of either their development potential (i.e., they were labeled "growth poles" by the World Bank) or of their high density of contract farming, as reported in the 2007 census of communes (Moser, 2008). In each region, the two communes with the highest density of contract farming were selected. The contracts in the data cover about a dozen crops. As discussed in Bellemare (2012), this diversity of crops and geographical areas ensures that our findings have more external validity than those of most other studies of contract farming, which focus at most on a handful of crops or on a more restricted geographical area.

Data collection was funded by the World Bank's Madagascar office for a study of the welfare effects of participation in contract farming. No pre-analysis plan was filed before the data were collected, but the primary goal of data collection was to study the effects of participation in contract farming on income, as in Bellemare (2012). Additionally, because of how the sample was constructed—in each commune, enumerators interviewed equal numbers of contract farming participants and nonparticipants—we follow the recommendations of Solon et al. (2015) and use sampling weights when computing descriptive statistics and when estimating the relationship between participation in contract farming and income variability. As is usual with sampling weights, those weights were computed using the sample and population proportion of contract farming participants and nonparticipants in a given commune, with the population proportion obtained from the 2007 commune census (Moser 2008).

Table 1 presents descriptive statistics for the variables we use in our empirical analysis ($n = 1,178$), as well as balance tests between the sub-sample of households that do not participate in contract farming ($n = 599$) and households that do ($n = 579$). Looking at the results of balance tests in the last column of Table 1, it is obvious that the variables retained for analysis are not orthogonal to a household's participation in contract farming, and so the empirical apparatus presented in section 3 is necessary if one is to attempt identifying the potential causal relationship flowing from participation in contract farming to income variability. For the remainder of this paper, to ensure the robustness of our findings, we look at three versions of our results: one that considers the variability of the household income level, one that considers the variability of household income (stated on a) per capita basis, and one that considers the variability of household income per adult equivalent (AE).[14]

[14]See Deaton (1997) for a discussion of why income per adult equivalent is a better measure of household welfare. For our analysis, we assign a weight of one to each individual between the ages of 15 and 65, a weight of 0.5 to each individual below the age of 15, and a weight of 0.75 to each individual older than 65.

Before proceeding, it is important to clarify what our income variable measures. From footnote 13 in Bellemare (2012):

> A household's total income includes (i) its income the sales of animals (cattle, pigs, sheep, goats, and poultry); (ii) its wages from various sources of labor (herding, agriculture, state, business, and other wages); (iii) its income from nonagricultural activities (crafts, trade, hunting and fishing, forestry, mining, pensions, transfers, and transportation); (iv) its income from leases (land, cattle, and equipment rentals), from sales of animal byproducts (milk and eggs), and from the sales of noncontracted crops; and (v) its income from contract farming.

5 Estimation Results and Discussion

We now turn to our empirical results. To do so, we begin with nonparametric results that show kernel density estimates of income variability for those households that participate in contract farming and for those that do not. Those nonparametric results do not control for observable confounding factors, much less unobservable ones, so we then turn to our parametric results, discussing in turn our core results and the mechanisms whereby participation in contract farming is likely to decrease income variability before moving on to PSM and DR results as well as other robustness checks.

5.1 Nonparametric Analysis

Before presenting kernel density estimates, we discuss the results of the ancillary regression in equation 20, whose squared residual we use to compute our heteroskedasticity measure of income variability. Table 2 presents the results of that regression. Here, note that we cluster standard errors at the community level, following the recommendations in Abadie et al. (2017). Given the small number of clusters, however, we will compute standard errors for our variable of interest (i.e., participation in contract farming) in our core regressions using the Wild bootstrap (Cameron et al. 2008).

As discussed, we use the square of the residual from equation 20 as our heteroskedasticity measure (i.e., H) of income variability. We plot kernel density estimates for H for households that participate in contract farming and households that do not in Figures 1 to 3 respectively for household income variability in levels, per capita (within the household), and per adult equivalent (also within the household). [15]

Figures 1 to 3 seems to suggest there is no systematic difference in income variability between the households that participate in contract farming and those that do not. The results in all those figures, however, only look at unconditional correlations between participation in contract farming and income variability. The parametric analyses we now turn to will help disen-

[15] The kernel density estimates in Figures 1 to 3 rely on nonstandardized versions of our proxies for income variability. For our regression and matching results, we standardize all three variables by first demeaning them and then dividing by their standard deviation so as to have their mean be centered on zero and their standard deviation equal to one.

tangle a potential causal relationship from this apparent lack of correlation.

5.2 Parametric Analyses

Recall that our H measure of income variability lends itself to two different tests, one a t-test of whether (unconditional) income variability is equal across households that participate in contract farming and households that do not, and one regression-based (i.e., conditional) test of whether income variability is the same across those two groups, holding RHS variables constant. A t-test that H is equal for participants and nonparticipants rejects the null at a significance level below 1 percent for the income of the household as well as for household income per capita and per adult equivalent within the household in favor of the alternative hypothesis that income variability is higher in the sub-sample of households that do not participate in contract farming.

For the regression-based approach, estimation results for H are shown in Table 3, where the first column shows results for the income of the household, whereas the next two columns respectively show results for household income per capita and per adult equivalent. The results in Table 3 shows that participation in contract farming is associated with a decrease in income variability of about 0.23 standard deviations, and that this association is significant at less than the 5 percent level in all cases. From these results, it looks as though participation in contract farming is associated with a decrease in income variability, providing evidence in support of Hypothesis

1 in Section 2.

5.2.1 Robustness Checks

As additional checks to ensure that the results in Table 3 are robust, we re-estimated median regression versions of the results in that table. Estimation results for those median regressions are shown in Appendix Table B1. In every case, results for those median regressions show once again that there is a negative and statistically significant relationship between participation in contract farming and income variability, although the estimated effects in those median regressions tend to be smaller in magnitude than those in the OLS regressions in Table 3 (i.e., a decrease in income variability of 0.08 vs. 0.20 standard deviation on average), but they also tend to have more statistical significance. Given that a median regression reduces the effect of outliers, it is no surprise that these coefficient estimates tend to be both smaller in magnitude and more precisely estimated than those of OLS regressions, as outliers can exaggerate estimated effects and make them more imprecise in the OLS case.

An anonymous reviewer also suggested we jointly estimate the income levels and variance equations in Tables 2 and 3 using the method in Western and Bloome (2009). In Table B2, we show the variance equations (income level equations not shown) estimated using that method. In all cases, the size of the relationship between participation in contract farming and income variability more than doubles relative to the other estimates in this paper.

5.2.2 Mechanisms

Regarding the mechanism whereby participation in contract farming reduces income variability, recall that Proposition 1 posited that contract farming insures growers against price risk via contracts in which they receive a fixed price. Hypotheses 2 and 3 were also concerned with this causal mechanism. In Table 4, we begin testing Hypothesis 2 by substituting the proportion of a household's plot that are under a fixed price contract for the treatment variable. In all three columns of Table 4, we find that the greater the proportion of a household's plots is under a fixed price contract, the lower the variability of that household's income; in each case, the relationship is significant at less than the 5 percent level. Specifically, a household whose plots would entirely be under fixed price contracts would see its income variability be about 0.25 standard deviations lower than that of a household whose plots would be entirely used to grow crops to be sold on spot markets or within contracts whose price is not fixed.

Our rejection of the null in this case provides support for Hypothesis 2, especially in light of the fact that no other variable is significantly associated with income variability. But we can go one step further in assessing whether fixed price contracts are a mechanism whereby participation in contract farming seems to provide partial insurance to grower households. In recent work, Acharya et al. (2016) develop a method that allows assessing whether a mediator (i.e., a variable that lies between the treatment and outcome variables on the causal path) is a mechanism whereby the treatment causes the outcome. In the limit, Acharya et al.'s method allows determining whether the mediator is the only mechanism, statistically speaking, whereby the treatment causes the outcome.

As in equation 23, let y be the outcome variable, and let D be the treatment variable. Moreover, let x^{Pre} denote control variables that are determined before the treatment is assigned, x^{Post} denote control variables that are determined after the treatment is assigned, and let M denote the presumed mechanism, or mediator variable; in our application, M is the proportion of a household's plots that are under a fixed price contract. Acharya et al.'s method then consists of the following steps:

1. Estimate $y = \alpha_3 + \beta_3^{\text{Pre}} x^{\text{Pre}} + \beta_3^{\text{Post}} x^{\text{Post}} + \gamma_3 D + \phi_3 M + \epsilon_3$.
2. Compute $\widetilde{y} = y - \widehat{\beta}_3^{\text{Post}} x^{\text{Post}} - \widehat{\phi}_3 M$.
3. Estimate $\widetilde{y} = \alpha_4 + \beta_4^{\text{Pre}} x^{\text{Pre}} + \gamma_4 D + \epsilon_4$.
4. The estimated parameter $\widehat{\gamma}_4$ is then the effect of the treatment once the mediator or mechanism M has been accounted for. In keeping with Acharya et al.'s terminology, we will hereafter refer to this effect as the "direct effect," in contrast with the "indirect effect" of treatment, which is the effect of treatment through the mechanism. If one fails to reject the null hypothesis $H_0 : \widehat{\gamma}_4 = 0$, one can then say that M is the only mechanism whereby the treatment D causes the outcome y.

In order to use Acharya et al.'s method, the only decision we need to make is to determine which of our control variables are pre- and post-treatment. In this case, we assume that (i) whether the household head is single, female, or a migrant, her age, education, and agricultural experience, whether she is a member of a farm organization, the number of days for which agricultural work is forbidden for her, which community she resides in, and her WTP for contract farming are pre-treatment, and (ii) the size of her household and her household's dependency ratio, the value of her assets and her working capital, and the size of her landholdings are post-treatment, since those latter variables could in theory change in response to treatment.[16]

The bottom part of Table 3 presents the direct effect coefficient (i.e., the effect of the dummy for participation in contract farming once that variable has been de-mediated by cleaning it out of the effect of the proportion of fixed price contracts) and the p-value on that coefficient for all three income measures. In each case, the direct effect is not only statistically insignificant, it is also very close to zero. This constitutes evidence that fixed price contracts are not only *a* mechanism whereby participation in contract farming is associated with a decrease in income variability, it is also evidence that it

[16] Again, because Chen (2008) and Chen and Risen (2010) show that choices reflect preferences instead of preferences reflecting choices, with the latter being based on flawed psychological studies, and so we include our respondents' preferences regarding contract farming in the set of pre-treatment variables.

is likely the *only* mechanism whereby this association holds in our data.[17,18] Thus, Hypothesis 3 is supported by our data.

Finally, looking at the correlation between income from contract farming and income from other sources, we find that that correlation is positive and significant at less than the 10 percent level between income from contract farming and income from nonfarm enterprises as well as income from agriculture, but that that correlation is not statistically significantly different from zero between income from contract farming and income from livestock as well as income from labor markets. Consequently, we can rule out the hypothesis that contract farming serves as partial insurance because income from contract farming is negatively related with income from other sources.[19]

[17] Given the conceptual setup wherein the proportion of fixed price contracts lies on the causal path between participation in contract farming and income variability, one might be tempted to try to identify the causal impact of participation in contract farming on income variability using Pearl's (2009) so-called front-door criterion (FDC). In order to apply the FDC, however, one would need to make the case that unobserved confounders that affect both participation in contract farming and income variability do not affect the proportion of a household's plots that are under a fixed-price contract. Given that unobserved confounders are unlikely to simultaneously (i) affect the outcome and the controls and (ii) not affect the mechanism, we do not look into the FDC.

[18] An anonymous reviewer asked about the difference between Acharya et al.'s (2016) two-stage method versus simply conditioning on the mediator in addition to the treatment variable and seeing whether the significance of the treatment is washed out. The latter method is a valid way of testing the mechanism in the case wherein all control variables are deemed pre-treatment. If some control variables are determined post-treatment, as in our application, simply including the mediator as a regressor may bias the coefficient estimate on the treatment. In this sense, the usual method of controlling for the mediator in addition to the treatment is nested within Acharya et al.'s method.

[19] An anonymous reviewer also wondered whether our findings might be explained by the fact that contracting households grow a different crop mix than non-contracting households. To address this, we looked at what were the most common crops in the data, and retained the top 10. For each of those crops, we regressed a dummy variable equal to one if that specific crop (e.g., maize) was grown on a given plot on a dummy variable equal to one if the household participated in contract farming (CF), applying sampling weights

5.2.3 Propensity Score Matching Results

Turning to our PSM results, Table 5 presents estimation results for equation 24, i.e., a probit aimed at predicting propensity scores. Similar probit results can be found in Bellemare (2012) and Bellemare and Novak (2017). Both those papers discuss the determinants of participation in contract farming, and since the probit results are, in this study, only interesting insofar as they allow predicting propensity scores, we encourage readers interested in those determinants to consult those two papers.

Appendix Figure B1 graphs histograms of propensity scores by participation regime. This common support graph shows that there is enough overlap in the propensity scores of participants and nonparticipants to yield reliable results. Additionally, Appendix Table B3 shows balance statistics for our matched sample. In no case do the means of the variables retained for analysis differ significantly between the treatment and control groups.

Our interest here is in estimating the ATE as well as the ATT and the ATU of participating in contract farming. Table 6 summarizes those estimates. Our estimates of the ATE of participating in contract farming on income variability are close to the ones we get from our regression analysis, seeing as to how they show a decrease of about 0.16 standard deviation in

---

and clustering at the same level as our core estimates. Results (not shown) indicate that there is some significance for the top two crops (i.e., maize and string beans), which are respectively 7.9 and 2.6 percentage points less likely to be grown on a household-plot if the household has a contract. That statistical significance, however, is rather low at less than the 10 percent level in both cases, and so this is unlikely to be a significant mechanisms behind our core finding.

the variability of income associated with participation in contract farming.

Though it is encouraging to see that our matching results confirm our regression results, what is even more interesting is the comparison between the ATT and the ATU. Intuitively, because farmers who would benefit the most should choose to participate in contract farming, one would expect the magnitude of the ATT to exceed that of the ATU. Here, however, the opposite result arises (i.e., the magnitude of the ATU exceeds that of the ATT) in all three cases. In other words, it looks as though considering income variability only, those households that do not participate in contract farming would benefit even more from participating in contract farming than those households that do participate, as the partial insurance derived from participation would be greater for nonparticipants than for participants. This is similar to Mishra et al.'s (2018b) results, wherein growers who did not participate in contract farming agreements would benefit from doing so.

5.2.4 Doubly Robust Weighted Regression Results

We now combine the regression approach in section 5.2.2 with the propensity scores in section 5.2.3 to generate doubly robust weighted regression estimates wherein propensity scores are used as regression weights.[20]

Table 7 is structured like Table 6 as it presents ATU, ATE, and ATT estimates for our variable of interest (i.e., participation in contract farming)

[20]These propensity scores are used as probability weights in addition to the sampling weights computed on the basis of the population and sample proportion of contract farming participants. Thus, the final weights used as part of this approach multiply propensity scores by the sampling weights.

and for our three outcome variables. Here, the rank ordering between the ATU, the ATE, and the ATT found by PSM remains, but the spread between these estimates is much tighter: Whether one looks at the ATU, the ATE, or the ATT, the estimated effect is generally that participation in contract farming leads to a decrease in income variability of about 0.23 standard deviations.

In sum, it looks as though participation in contract farming can be an effective partial insurance mechanism for households in rural Madagascar, with estimated ATEs of about -0.22 to standard deviation in our regression results, of about -0.17 standard deviation in our matching results, and of about -0.23 standard deviations in our doubly robust weighted regression results. Moreover, our investigation of the mechanisms whereby contract farming can serve as partial insurance support Proposition 1, according to which fixed-price contracts are the main mechanism whereby this happens. Finally, it looks as though those households that would benefit the most from participating in contract farming when it comes to reductions in income variability are those who do not participate, as the ATU always exceeds the ATT.

6 Summary and Concluding Remarks

We have looked at whether participation in contract farming can serve as partial insurance for rural households, i.e., whether participating households experience lower levels of income variability. To do so, we have used the results of a framed field experiment aimed at eliciting WTP for participation in a hypothetical contract farming agreement that would raise the respondent's income level by 10 percent in an effort to exogenize actual participation in contract farming—our treatment variable—in a selection-on-observables design. Given that our design relies on the same assumption which makes propensity score matching credible, we supplement our core regression approach with a matching approach.

Both approaches lead to comparable estimates of the average treatment effect: in most cases, participation in contract farming is associated with a 0.20-standard deviation decrease in income variability, and so contract farming appears to offer participating households a certain degree of partial insurance. Looking at the mechanism behind our main result, we use an empirical method newly developed by Acharya et al. (2016) and find that, in line with our theoretical prediction, fixed price contracts are not only a mechanism whereby participation in contract farming provides partial insurance, those same fixed price contracts appear to be the *only* mechanism whereby this happens.

From a welfare perspective, though it is difficult to assess the economic—not statistical—significance of our estimates of the effect of participation in contract farming on income variability, given that there is no accepted way of measuring income variability and that we rely on a standardized measure of it, we can speculate about the welfare aspects of our findings. If one is to assume that our findings are economically significant and that the average respondent in our sample is risk-averse, then it is clear that participation in contract farming improves the welfare of the households involved via second-order (i.e., expected-utility) effects. This militates in favor of encouraging contracts wherein the price for the contracted crop is fixed, thereby letting the processor bear all of the price risk. This is in line with empirical findings in Bellemare et al. (2013), Bellemare et al. (2020), and Lee (2020) for net producers and their view of output price risk.

Perhaps more importantly for development policy, our findings indicate that the usual intuitive ordering of average treatment effects between the treated and the untreated is reversed. That is, the counterfactual analysis our matching approach provides shows that those households that do not participate in contract farming would benefit from participating even more than those households that do participate—the untreated would receive a higher degree of partial insurance than the treated. Though it is impossible to determine why those households that would benefit the most do not participate, one can speculate that it is likely because households are more likely to choose to participate on the basis of an expected higher income level rather than of a lower expected income variability.

Our analysis is not without its limitations, and we wish to note three important limitations of our work. First, given our research design, our results cannot be argued to be causal, though we claim that we control for the most important sources of statistical endogeneity with our framed field experiment aimed at eliciting respondent WTP—and thus marginal utility—for contract farming. Second, in the absence of longitudinal data, our dependent variables are only proxies for income variability. Third, income variability is but one measure of risk; as a reviewer encouraged us to note, citing Andersson et al. (2015) and Ochieng et al. (2017), contract farming may be associated with certain risks over time that cannot be analyzed with the cross-section data at hand. Such risks may include unexpected changes in contract clauses and conditions (e.g., stricter quality requirements), loss of the support (technical or financial) previously received, and so on. We leave the use of better research designs combined with longitudinal data to future research.

References

[1] Abadie, Alberto, Susan Athey, Guido Imbens, and Jeffrey Wooldridge (2017), "When Should You Adjust Standard Errors for Clustering?," Working Paper, Massachusetts Institute of Technology.

[2] Acharya, Avidit, Matthew Blackwell, and Maya Sen (2016), "Explaining Causal Findings Without Bias: Detecting and Assessing Direct Effects," *American Political Science Review* 110(3): 512–529.

[3] Andersson, Camilla I., Christine G.K. Chege, Elizaphan James Rao, and Matin Qaim (2015), "Following Up on Smallholder Farmers and Supermarkets in Kenya," *American Journal of Agricultural Economics* 97(4): 1247-1266.

[4] Arouna, Aminou, Jeffrey D. Michler, and Jourdain C. Lokossou (2019), "Contract Farming and Rural Transformation: Evidence from a Field Experiment in Benin," NBER Working Paper #25665.

[5] Baron, David (1970), "Price Uncertainty, Utility, and Industry Equilibrium in Pure Competition," *International Economic Review* 11: 463—480.

[6] Barrett, Christopher B. (1996), "On Price Risk and the Inverse Farm Size-Productivity Relationship," *Journal of Development Economics*, 51: 193—215.

[7] Begum, Ismat A. (2006), "Potentiality of Vertically Integrated Poultry Contract Farming System in Bangladesh: Profitability and Productivity Analysis," *Bangladesh Journal of Agricultural Economics* XXIX: 107–117.

[8] Bellemare, Marc F. (2012), "As You Sow, So Shall You Reap: The Welfare Impacts of Contract Farming," *World Development* 40(7): 1418–1434.

[9] Bellemare, Marc F., Christopher B. Barrett, and David R. Just (2013), "The Welfare Impacts of Commodity Price Volatility: Evidence from Rural Ethiopia," *American Journal of Agricultural Economics* 95(4): 877–899.

[10] Bellemare, Marc F., and Jeffrey R. Bloem (2018), "Does Contract Farming Improve Welfare? A Review," *World Development* forthcoming.

[11] Bellemare, Marc F., Yu Na Lee, and David R. Just (2020), "Producer Attitudes Toward Output Price Risk: Experimental Evidence from the Lab and from the Field," *American Journal of Agricultural Economics* 102(3): 806–825.

[12] Bellemare, Marc F., and Sunghun Lim (2018), "In All Shapes and Colors: Varieties of Contract Farming," *Applied Economic Perspectives and Policy*, forthcoming.

[13] Bellemare, Marc F., and Lindsey Novak (2017), "Contract Farming and Food Security," *American Journal of Agricultural Economics* 99 (2): 357–378.

[14] Bellemare, Marc F., and Casey J. Wichman (2020), "Elasticities and the Inverse Hyperbolic Sine Transformation," *Oxford Bulletin of Economics and Statistics* 82(1): 50-61.

[15] Bijman, Jos (2008), "Contract farming in developing countries: An overview," *Working Paper*, Wageningen University.

[16] Bolwig, Simon, Peter Gibbon, and Sam Jones (2009), "The Economics of Smallholder Organic Contract Farming in Sub-Saharan Africa," *World Development* 37(6): 1094–1104.

[17] Briones, Roelhano M. (2015), "Small Farmers in High-Value Chains: Binding or Relaxing Constraints to Inclusive Growth," *World Development* 72: 43–52.

[18] Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller (2008), "Bootstrap-Based Improvements for Inference with Clustered Errors," *Review of Economics and Statistics* 90(3): 414-427.

[19] Chen, M. Keith (2008), "Rationalization and Cognitive Dissonance: Do Choices Affect or Reflect Preferences?," Working Paper, Cowles Foundation for Research in Economics, Yale University.

[20] Chen, M. Keith, and Jane L. Risen (2010), "How Choice Affects and Reflects Preferences: Revisiting the Free-Choice Paradigm," *Journal of Personality and Social Psychology* 99(4): 573-594.

[21] Deaton, Angus S. (1997), *The Analysis of Household Surveys: Microeconometric Analysis for Development Policy* Baltimore, MD, Johns Hopkins University Press.

[22] Dedehouanou, Senakpon F.A., Johan Swinnen, and Miet Maertens (2013), "Does Contracting Make Farmers Happy? Evidence from Senegal," *Review of Income and Wealth* 139(S1): S138–S160.

[23] Dercon, S. and P. Krishnan (1998), "Changes in Poverty in Rural Ethiopia 1989-1995: Measurement, Robustness Tests and Decomposition," Working Paper, Centre for the Study of African Economies, Oxford University.

[24] Du, Xiaoxue, Liang Lu, Thomas Reardon, and David Zilberman (2016), "Economics of Agricultural Supply Chain Design: A Portfolio Selection Approach," *American Journal of Agricultural Economics* 98(5): 1377-1388.

[25] Engle, Robert F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica* 50(4): 987-1007.

[26] Foster, Andrew D., and Mark R. Rosenzweig (1995), "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *Journal of Political Economy* 103(6): 1176-1209.

[27] Glover, David (1990), "Contract Farming and Outgrower Schemes in East and Southern Africa," *Journal of Agricultural Economics* 41(3): 303–315.

[28] Goldberger, Arthur S. (1991), *A Course in Econometrics*, Cambridge, MA: Harvard University Press.

[29] Grosh, Barbara (1994), "Contract Farming in Africa: An Application of the New Institutional Economics," *Journal of African Economies* 3(2): 231–261.

[30] Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart (2007), "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference," *Political Analysis* 15(3): 199-236.

[31] Imbens, Guido (2015), "Matching Methods in Practice: Three Examples," *Journal of Human Resources* 50(2):373–419.

[32] Jones, Sam and Peter Gibbon (2011), "Developing Agricultural Markets in Sub-Saharan Africa: Organic Cocoa in Rural Uganda, "*Journal of Development Studies* 47(10):1595–1618.

[33] Kumar, Jagdish and Prakash K. Kumar (2008), "Contract farming: Problems, prospects and its effect on income and employment, "*Agricultural Economics Research Review* 21(2):243–250.

[34] Lee, Yu Na (2020), "Does an Aversion to Price Risk Drive Migration? Evidence from Rural Ethiopia," Working Paper, University of Guelph.

[35] List, John A. (2011), "Why Economists Should Conduct Field Experiments and 14 Tips for Pulling One Off, "*The Journal of Economic Perspectives* 25(3): 3–15.

[36] Lu, Liang, and Thomas Reardon (2018), "An Economic Model of the Evolution of Food Retail and Supply Chains from Traditional Shops to Supermarkets to E-Commerce," *American Journal of Agricultura Economics* forthcoming.

[37] Lu, Liang, Thomas Reardon, and David Zilberman (2016), "Supply Chain Design and Adoption of Indivisible Technology," *American Journal of Agricultural Economics* 98(5): 1419-1431.

[38] Maertens, Miet, and Johan F.M. Swinnen (2009), "Trade, Standards, and Poverty: Evidence from Senegal," *World Development* 37(1): 161–178.

[39] Michelson, Hope C., Thomas Reardon, and Francisco Perez (2012), "Small Farmers and Big Retail: Trade-Offs of Supplying Supermarkets in Nicaragua," *World Development* 40(2): 342-354.

[40] Minten, Bart, Lalaina Randrianarison, and Johan F.M. Swinnen (2009), "Global Retail Chains and Poor Farmers: Evidence from Madagascar," *World Development* 37(11): 1728–1741.

[41] Mishra, Ashok K., Saleem Shaik, Aditya Khanal, and Subir Bairagi (2018a), "Contract Farming and Technical Efficiency: Evidence from Low-Value and High-Value crops in Nepal," *Agribusiness* 34(2): 426-440.

[42] Mishra, Ashok K., Anjani Kumar, Pramod K. Joshi, Alwin D'Souza, and Gaurav Tripathi (2018b), "How Can Organic Rice Be a Boon to Smallholders? Evidence from Contract Farming in India," *Food Policy* 75: 147-157.

[43] Miyata, Sachiko, Nicholas Minot, and Dinghuan Hu (2009), "Impact of Contract Farming on Income: Linking Small Farmers, Packers, and Supermarkets in China," *World Development* 37(11): 1781–1790.

[44] Montalbano, Pierluigi, Rebecca Pietrelli, and Luca Salvatici (2018), "Participation in the Market Chain and Food Security: The Case of the Ugandan Maize Farmers," *Food Policy* 76: 81-98.

[45] Morgan, Stephen L., and Christopher Winship (2015), *Counterfactuals and Causal Inference*, Cambridge: Cambridge University Press.

[46] Moser, Christine M. (2008) "Report on the 2007 Commune Census," *Working Paper* Western Michigan University.

[47] Mwambi, Mercy Maiwa, Judith. Oduol, Patience Mshenga, and Mwanarusi Saidi (2016) "Does contract farming improve smallholder income? The case of avocado farmers in Kenya," *Journal of Agribusiness in Developing and Emerging Economies* 6(1): 2–20.

[48] Naryananan, Sudha (2014), "Profits from Participation in High-Value Agriculture: Evidence of Heterogeneous Benefits in Contract Farming Schemes in Southern India," *Food Policy* 44: 142–157.

[49] Ochieng, Dennis O., Prakashan Chellattan Veettil, , and Matin Qaim (2017), "Farmers' preferences for supermarket contracts in Kenya," *Food Policy* 68, 100-111.

[50] Pearl, Judea (2009), *Causality*, 2nd ed., Cambridge: Cambridge University Press.

[51] Raynolds, Laura T. (2002), "Wages for Wives: Renegotiating Gender and Production Relations in Contract Farming in the Dominican Republic," *World Development* 30(5): 783–798.

[52] Ruud, Jørgen (1960), *Taboo: A Study of Malagasy Customs and Beliefs*, Oslo: Oslo University Press.

[53] Sandmo, Agnar (1971), "On the Theory of the Competitive Firm under Price Uncertainty," *American Economic Review* 59: 138—148.

[54] Sharma, Nivedita (2008), "Contract Farming Practice in Indian Punjab: Farmers' Perspective," *International Journal of Food and Agricultural Economics* 2(1): 65–76.

[55] Simmons, Phil, Paul Winters, and Ian Patrick (2005), "An Analysis of Contract Farming in East Java, Bali, and Lombok, Indonesia," *Agricultural Economics* 33(s3): 513–525.

[56] Singh, Sukhpal (2002), "Contracting Out Solutions: Political Economy of Contract Farming in the Indian Punjab," *World Development* 30(9): 1621–1638.

[57] Solon, Gary, Steven J. Haider and Jeffrey M. Wooldridge (2015), "What Are We Weighting For?," *The Journal of Human Resources* 50(2): 301–316.

[58] Stiglitz, Joseph E. (1974), "Incentives and Risk Sharing in Sharecropping," *The Review of Economic Studies* 41(2): 219–255.

[59] Trifković, Neda (2016), "Vertical coordination and farm performance: evidence from the catfish sector in Vietnam," *Agricultural Economics* 47(5): 547-–557.

[60] Wainaina, Priscilla W., Julius J. Okello, and Jonathan M. Nzuma (2014) "Blessing or Evil? Contract Farming, Smallholder Poultry Production and Household Welfare in Kenya," *Quarterly Journal of International Agriculture* 53(4): 319–340.

[61] Wang, Holly H., Yanbing Wang, Michael S. Delgado (2014) "The transition to modern agriculture: Contract farming in developing countries," *American Journal of Agricultural Economics* 96(5): 1257–1271.

[62] Warning, Matthew, and Nigel Key (2002), "The Social Performance and Distributional Consequences of Contract Farming: An Equilibrium Analysis of the Arachide de bouche Program in Senegal," *World Development* 30(2): 255–263.

[63] Western, Bruce, and Deirdre Bloome (2009), "Variance Function Regressions for Studying Inequality," *Sociological Methodology* 39(1): 293-326.

[64] Zilberman, David, Liang Lu, and Thomas Reardon (2017), "Innovation-Induced Food Supply Chain Design," *Food Policy* forthcoming.

Appendix A

In this appendix, we investigate the validity of our proxy measure of income variability by comparing it with longitudinal measures of household income variability. That is, we investigate the assumption that the squared absolute distance between a household's income realization and the income predicted based on the control variables, $H_i = \hat{\epsilon}_i^2$, is representative of the same household's longitudinal income variation.

Recall that our data is cross-sectional, so we cannot verify this assumption with our data. Instead we use data from two well-known data sets, viz. the Tanzania Living Standards Measurement Study (LSMS) and the Ethiopian Rural Household Survey (ERHS). Both of these surveys are longitudinal and interview agrarian households about their income and income generating activities.

A.1 Tanzania Living Standards Measurement Study

The Tanzania LSMS includes three waves of data in which the same households were interviewed in 2009, 2011, and 2013. We first estimate the intra-household income variation, V_i, using the three waves of data, where

$$V_i = \sum_{t=1}^{3} \left(y_{it} - \frac{1}{3} \sum_{t=1}^{3} (y_{it}) \right)^2$$

and y_{it} is the inverse hyperbolic sine of household i's income at the time of survey wave $t \in \{1, 2, 3\}$. We compare V_i to our conditional heteroskedasticity measure $H_i = \hat{\epsilon}_i^2$ where $\hat{\epsilon}_i$ is found by estimating equation (20)

$$IHS(y_i) = \alpha_0 + \underline{\beta}_0 \underline{x}_i + \gamma_0 D_i + \epsilon_{0i},$$

where $IHS(\cdot)$ denotes the inverse hyperbolic sine transformation, used here to account for those cases where a household reports no income in a given time period (Bellemare and Wichman 2020). In estimating equation (20), we use the same explanatory variables used in our primary analysis wherever possible. Recall that D is an indicator variable for whether a household participates in contract farming, and $\underline{x}_i$ is a vector of household-specific control variables: household head marital status, household head gender, household head age, household head education level, the size of the household's landholdings, size of the household, and the household's dependency ratio. The LSMS does not ask respondents if they are members of a farm organization but does ask village representatives if there is a farm organization active in the village. We include this binary variable in the regression. The LSMS includes a measure of the value of the household's assets in the third wave of data collection only. Finally, the LSMS does not include a measure of the years of agricultural experience of the household head, the number of days that farming is forbidden, nor the value of the household's working capital. Table A1 displays descriptive statistics for the included variables. In these analyses, we restrict the sample to households located in a rural area in order to make the sample more comparable to our Madagascar sample.

In order to be included in this analysis, households must have been surveyed and reported their income in all three survey waves. The variation in sample size is due to missing data in one or more explanatory variables.

Table A2 displays the correlation coefficients between our measure of income variation using cross-sectional data, H_i, and the intrahousehold income variation using the longitudinal survey, V_i.

Table A2 shows that, among rural Tanzanian households, our measure of income variation, while not perfectly correlated, is highly predictive of within-household longitudinal income variation. The results from wave 1 are likely the most similar to the results we would obtain from our Madagascar data as the later survey waves suffer from missing data. Because our Madagascar data are cross-sectional, our set of respondents is likely to be more representative of the population, similar to wave 1 of the LSMS compared to later waves.

We can also determine for which subgroups our proxy is most and least relevant. Table A3 displays the correlation coefficients between within-household longitudinal income variation, V_i, and our income variation proxies by subgroups using data from wave 1 of the LSMS.

Table A3 shows that our cross-sectional measure of income variation is a better proxy for households with (1) at least seven members, (2) a dependency ratio above 0.5, a household head who is (3) married, (4) male, (5) migrated to their current district, (6) younger than 65, (7) has some education, and (8) has landholdings less than the median amount in the sample.

Table A1: Descriptive Statistics from Tanzania LSMS Data

	Wave 1	Wave 2	Wave 3
Household Income (USD 2016 PPP)	5,505.761	5,820.824	5,685.429
	(20,005.822)	(13,926.315)	(18,239.828)
Contract Farming Participant	0.011	0.017	0.020
	(0.106)	(0.130)	(0.141)
Household Size (Individuals)	5.408	5.751	5.767
	(2.956)	(3.233)	(3.251)
Dependency Ratio (Proportion)	0.482	0.487	0.486
	(0.236)	(0.227)	(0.238)
Household Head Single (Dummy)	0.229	0.213	0.242
	(0.420)	(0.409)	(0.428)
Household Head Female (Dummy)	0.237	0.225	0.242
	(0.425)	(0.418)	(0.429)
Household Head Migrant (Dummy)	0.283	0.302	0.415
	(0.451)	(0.459)	(0.493)
Household Head Age (Years)	47.184	49.748	51.001
	(15.609)	(15.569)	(15.504)
Household Head Education			
Never Attended (Dummy)	0.287	0.305	0.291
	(0.452)	(0.461)	(0.455)
Attended Primary (Dummy)	0.630	0.604	0.613
	(0.483)	(0.489)	(0.487)
Attended Secondary (Dummy)	0.076	0.085	0.090
	(0.264)	(0.279)	(0.287)
Attended Tertiary (Dummy)	0.007	0.005	0.005
	(0.083)	(0.071)	(0.070)
Farmer Coop in Community (Dummy)	0.517	0.524	0.487
	(0.500)	(0.500)	(0.500)
Household Landholdings (Hectares)	1.997	2.185	2.256
	(6.826)	(3.766)	(5.527)
Household Assets (USD 2016 PPP)	-	-	5,191.376
			(13,602.547)
Observations	1,864	1,575	1,630

Standard deviations in parentheses.

Table A2: Correlation coefficients between V_i and our cross-sectional measures of income variation: LSMS data

	Wave 1	Wave 2	Wave 3
H_i	0.640	0.541	0.467

Table A3: Correlation coefficients between V_i and our cross-sectional measures of income variation for sub-samples using data from Wave 1

	HH Size		Dep. Ratio		Marital Status	
	≤ 6	> 6	≤ 0.5	> 0.5	Single	Married
H_i	0.631	0.670	0.595	0.672	0.454	0.668
	Sex		Migrant		Age	
	Female	Male	Yes	No	≤ 65	> 65
H_i	0.448	0.693	0.675	0.628	0.657	0.452
	Education		Landholdings		Contract Farmer	
	None	Some	$\leq$ Median	$>$ Median	Yes	No
H_i	0.403	0.685	0.678	0.579	0.895	0.643

The results for contract farming participation are less clear which is likely due to the small sample size of contract farmers in this sample.

A.2 Ethiopian Rural Household Surveys

The Ethiopian Rural Household Survey (ERHS) is a household panel that covers rural villages in Ethiopia. There are eight survey rounds, namely, 1989, 1994a, 1994b, 1995, 1997, 1999, 2004, and 2009. In this appendix, we focus on 1994a, 1994b, 1995, 1997 rounds of the survey, given that the questionnaires are highly comparable and attrition rate was particularly low across these rounds. (Dercon and Krishnan, 1998).

As the ERHS is a panel data set, we can estimate the household income variation over time, V_i, using the four rounds of survey data, where

$$V_i = \sum_{t=1}^{4} \left(y_{it} - \frac{1}{4} \sum_{t=1}^{4} (y_{it}) \right)^2$$

and y_{it} is the inverse hyperbolic sine of household i's income at the time of survey round $t \in \{1, 2, 3, 4\}$.

Our proxy for V_i is the conditional heteroskedasticity measure $H_i = \hat{\epsilon}_i^2$ where $\hat{\epsilon}_i$ is obtained by estimating equation (20)

$$IHS(y_i) = \alpha_0 + \underline{\beta}_0 \underline{x}_i + \gamma_0 D_i + \epsilon_{0i},$$

where D is an indicator for participation in contract farming, and $\underline{x}_i$ is a vector of household-level control variables. Unlike our Madagascar data and the Tanzanian LSMS, the ERHS does not include a variable on contract farming. Thus, the variable D is not included here. As in the case of the LSMS, the ERHS does not include a variable that measures whether a household is a member of a farm organization, but it includes a village-level indicator variable for an existence of a functioning producer cooperative in the village only in the 1997 survey round. Thus we include this variable. Unlike our Madagascar data, the ERHS does not include a variable on the agricultural experience of the household head. The ERHS, however, includes a measure of whether the occupation of the household head is farmer, which we include in place of the years of agricultural experience whenever available. Lastly, the ERHS does not include a household's working capital, the number of taboo days, or a measure of household asset. Instead, land and livestock holdings are included as proxies for wealth. All other household-level control variables in the vector $\underline{x}_i$ in the Madagascar sample, where available, are included, except for some survey rounds wherein comparable and reliable measures are not available.

Table A4 displays descriptive statistics for the household-level control variables included to estimate the equation above. The ERHS data is collected from rural villages of Ethiopia, which makes it comparable to our Madagascar data. As in the LSMS data, there exist some variation in the number of observations across the four rounds of the ERHS, which is due to the missing values of some variables.

Table A5 shows the correlation coefficients between our proxy measure of income variation, H_i, and the intrahousehold income variation using the longitudinal survey, V_i, separately for each survey round. The results demonstrate that among the sample Ethiopian rural households, our cross-sectional proxy measure of income variability is positively and significantly correlated with longitudinal, within-household income variation, with bivariate correlation coefficients ranging from 0.312 to 0.379. In all the survey rounds, the correlation coefficients are statistically significant at less than 1% level of significance.

As in the case of the LSMS data, the first (1994a) survey round suffers the least from missing data. Thus, we further examine which subgroups our proxies are most and least relevant in case of the ERHS data set, using the 1994a round. Table A6 shows the correlation coefficients between V_i and our proxy for V_i by subgroups using data from the 1994a survey round. Table

Table A4: Descriptive Statistics from ERHS Data

VARIABLES	(1) 1994a	(2) 1994b	(3) 1995	(4) 1997
Household Income (Birr)	656.768	664.572	601.170	962.304
	(957.626)	(2,583.355)	(1,220.078)	(2,225.584)
Household Size (Individuals)	6.091	6.081	5.999	5.780
	(3.038)	(3.034)	(3.000)	(2.723)
Dependency Ratio (Proportion)	0.551	0.536	0.523	0.482
	(0.192)	(0.192)	(0.192)	(0.196)
Household Head Single (Dummy)	0.250	–	0.230	0.248
	(0.433)		(0.421)	(0.432)
Household Head Female (Dummy)	0.216	–	0.217	0.242
	(0.412)		(0.412)	(0.429)
Household Head Migrant (Dummy)	0.017	0.004	0.004	0.002
	(0.131)	(0.067)	(0.061)	(0.047)
Household Head Farmer (Dummy)	0.705	–	–	0.686
	(0.456)			(0.464)
Household Head Age (Years)	46.696	47.265	47.505	48.477
	(15.886)	(15.882)	(15.897)	(15.279)
Farmer Coop in Community (Dummy)	–	–	–	0.466
				(0.499)
Household Landholdings (Hectares)	1.170	1.188	1.284	1.311
	(1.385)	(1.396)	(1.504)	(1.423)
Livestock Holdings (TLU)	2.461	2.460	2.415	3.082
	(3.260)	(3.292)	(3.274)	(3.551)
HH Head Never Attended School (Dummy)	0.725	–	–	–
	(0.447)			
HH Head Attended Primary (Dummy)	0.191	–	–	–
	(0.393)			
HH Head Attended Secondary (Dummy)	0.052	–	–	–
	(0.223)			
HH Head Attended Tertiary (Dummy)	0.002	–	–	–
	(0.047)			
Observations	1,375	1,348	1,350	1,357

Table A5: Correlation coefficients between V_i and our cross-sectional measures of income variation: ERHS data

	1994a	1994b	1995	1997
H_i	0.370	0.377	0.379	0.312

Note: All the correlation coefficients are statistically significant at less than 1% level of significance, each with p-value = 0.0000.

A6 shows that our cross-sectional proxy measures of income variation work better for households with (1) six or less members, (2) a dependency ratio less than or equal to 0.5, a household head who (3) is single, (4) is male, (5) did not move out of the current village, (6) is older than 65, (7) has no education, (8) has landholdings less than the median in the sample, and (9) has less than the median amount of assets.

Table A6: Correlation coefficients between V_i and our cross-sectional proxies of income variation for sub-samples using data from the 1994a Survey Round

	HH Size		Dep. Ratio		Marital Status	
	≤ 6	> 6	≤ 0.5	> 0.5	Married	Single
H_i	0.395	0.334	0.378	0.365	0.310	0.391
	Sex		**Migrant**		**Age**	
	Female	Male	Yes	No	≤ 65	> 65
H_i	0.333	0.380	0.0600	0.371	0.357	0.411
	Education		**Landholdings**		**Assets**	
	None	Some	Small	Large	Small	Large
H_i	0.392	0.322	0.408	0.340	0.400	0.351

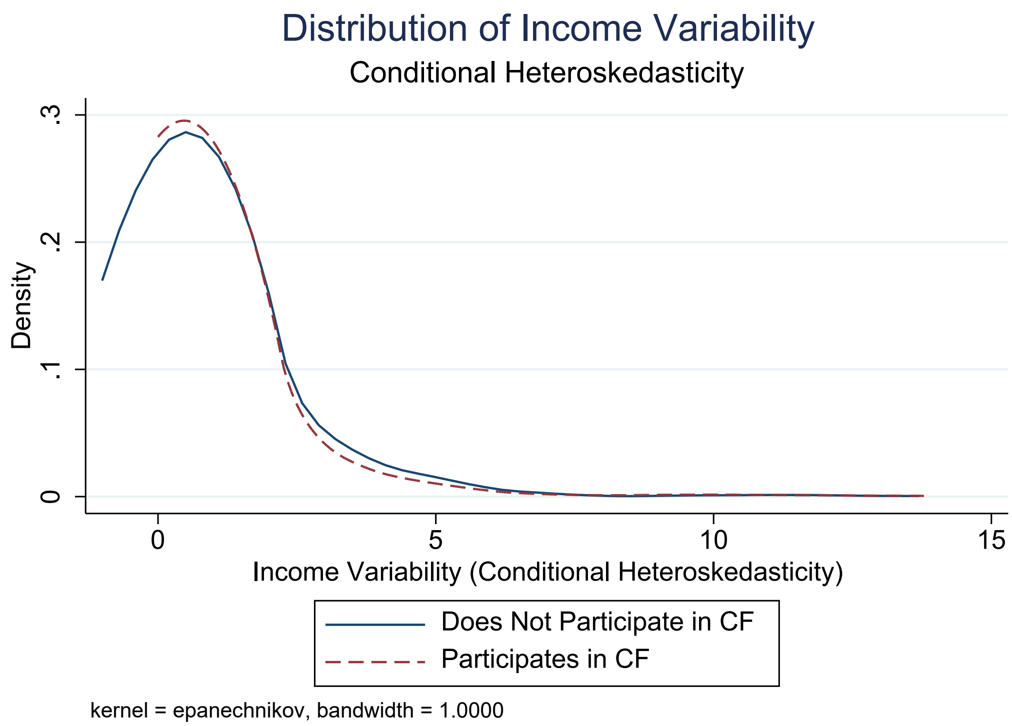

Figure 1. Kernel Density Estimates of Income Variability – Conditional Heteroskedasticity.

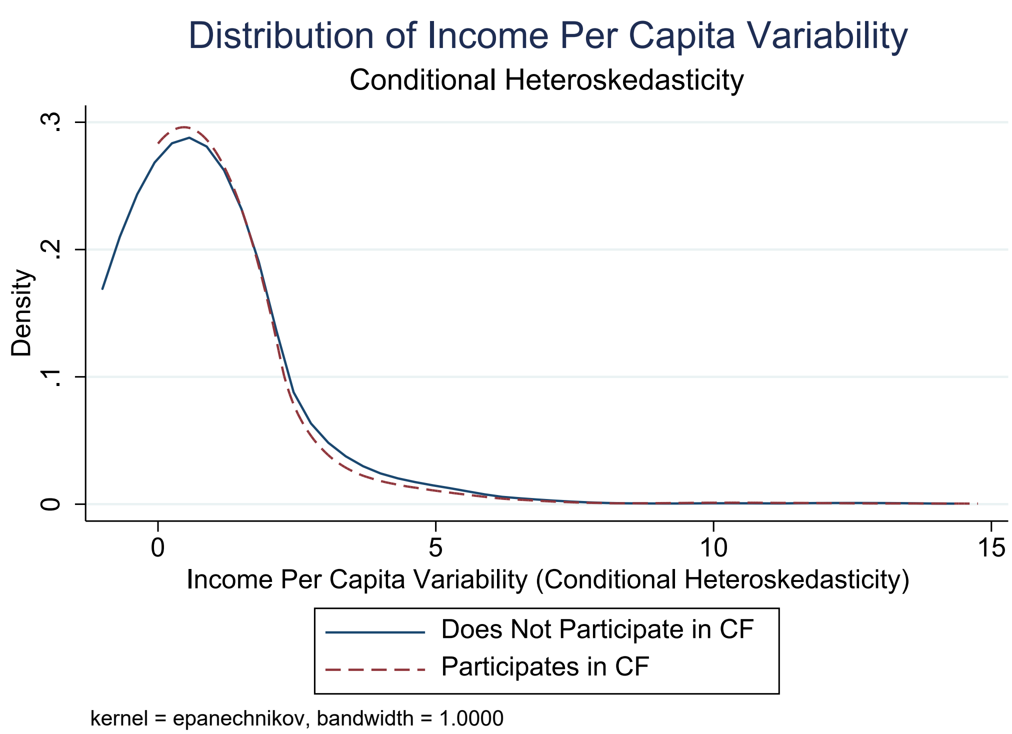

Figure 2. Kernel Density Estimates of Income Per Capita Variability – Conditional Heteroskedasticity.

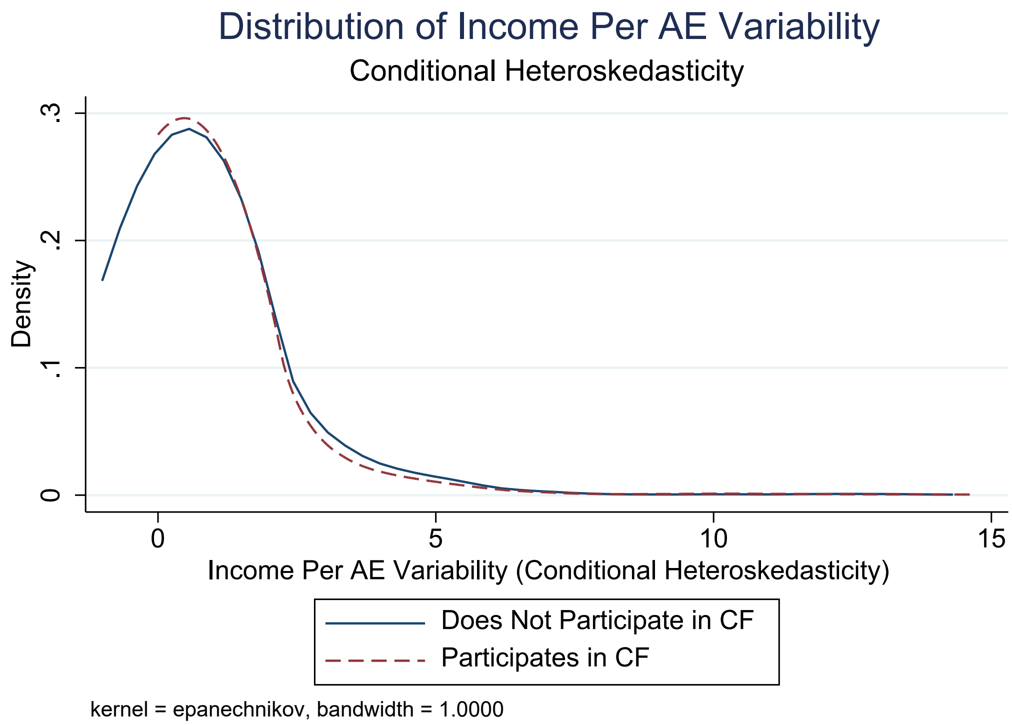

Figure 3. Kernel Density Estimates of Income Per AE Variability – Conditional Heteroskedasticity.

Table 1. Descriptive Statistics and Balance Tests (n=1,178)

Variables	Contract Farming [a] No	Contract Farming [a] Yes	Test of Difference [b]
Household Income	14.843	24.255	***
(Ariary)	(1.198)	(2.762)	
Household Income Per Capita	3.072	4.463	***
(Ariary)	(0.239)	(0.413)	
Household Income Per Adult Equivalent	3.802	5.535	***
(Ariary)	(0.294)	(0.471)	
Household Size	5.452	5.692	**
(Individuals)	(0.108)	(0.104)	
Dependency Ratio	0.452	0.446	
(Proportion)	(0.012)	(0.010)	
Household Head Single	0.158	0.089	***
(Dummy)	(0.017)	(0.014)	
Household Head Female	0.119	0.057	***
(Dummy)	(0.016)	(0.011)	
Household Head Migrant	0.124	0.125	
(Dummy)	(0.015)	(0.015)	
Household Head Age	44.428	42.110	***
(Years)	(0.652)	(0.554)	
Household Head Education (Years)	5.650	5.715	
(Years)	(0.154)	(0.147)	
Household Head Experience (Years)	21.074	20.165	
(Years)	(0.653)	(0.566)	
Household Head Member of a Farm Organization (Dummy)	0.149	0.296	***
	(0.017)	(0.022)	
Household Head Taboo Days[c]	23.968	20.427	*
(Days)	(1.684)	(1.424)	
Household Working Capital	2.872	6.021	***
(Ariary)	(0.380)	(0.973)	
Household Assets	11.672	16.277	***
(Ariary)	(1.099)	(1.359)	
Household Landholdings	113.438	177.956	***
(Ares)	(8.982)	(18.146)	
Yes to $12.50 Investment	0.672	0.800	***
(Dummy)	(0.022)	(0.019)	
Yes to $25.00 Investment	0.543	0.665	***
(Dummy)	(0.023)	(0.022)	
Yes to $37.50 Investment	0.371	0.480	***
(Dummy)	(0.022)	(0.023)	
Yes to $50.00 Investment	0.229	0.307	***
(Dummy)	(0.019)	(0.021)	

Yes to \$62.50 Investment	0.112	0.157
(Dummy)	(0.014)	(0.017)
Yes to \$75.00 Investment	0.047	0.085
(Dummy)	(0.009)	(0.013)
Observations	599	579

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

[a] District dummies omitted for brevity. Conditional means calculated using sampling weights.

[b] Tests of differences in conditional means do not use sampling weights.

[c] The Malagasy observe a complex system of taboos (known as *fady* in the local language) and interdictions, one of which is the interdiction to do agricultural work on certain days of the week, which we use as a control variable in the empirical analysis in this paper. For the multiplicity of taboos observed by the Malagasy, see Ruud (1960).

Table 2. Ordinary Least Squares Estimation Results for an Ancillary Income Regressions

Variables	(1) Income	(2) Income Per Capita	(3) Income Per AE
Dependent Variable: Log of Income			
Contract Farming	0.354***	0.343***	0.345***
	(0.052)	(0.051)	(0.051)
Household Size	0.051**	-0.129***	-0.122***
	(0.019)	(0.018)	(0.018)
Dependency Ratio	-0.129	-0.336*	0.204
	(0.142)	(0.147)	(0.132)
Single	-0.165	0.044	0.046
	(0.105)	(0.089)	(0.087)
Female	-0.329*	-0.466**	-0.455**
	(0.147)	(0.116)	(0.116)
Migrant	0.028	0.032	0.038
	(0.073)	(0.069)	(0.072)
Age	0.005	0.005	0.004
	(0.004)	(0.003)	(0.003)
Education	0.069***	0.071***	0.071***
	(0.009)	(0.009)	(0.009)
Agricultural Experience	-0.001	-0.001	-0.001
	(0.005)	(0.004)	(0.004)
Member of a Farm Organization	0.168***	0.167***	0.172***
	(0.028)	(0.033)	(0.034)
Taboo Days	0.001	0.000	0.000
	(0.001)	(0.001)	(0.001)
Working Capital	0.007**	0.008**	0.008**
	(0.003)	(0.002)	(0.002)
Assets	0.007***	0.007***	0.007***
	(0.002)	(0.002)	(0.002)
Landholdings	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Yes to $12.50 Investment	0.204	0.181	0.183
	(0.144)	(0.142)	(0.143)
Yes to $25.00 Investment	-0.113	-0.118	-0.117
	(0.193)	(0.191)	(0.192)
Yes to $37.50 Investment	0.020	0.048	0.048
	(0.081)	(0.071)	(0.069)
Yes to $50.00 Investment	-0.058	-0.075	-0.076

	(0.108)	(0.110)	(0.109)
Yes to $62.50 Investment	0.198*	0.197	0.200
	(0.089)	(0.102)	(0.100)
Yes to $75.00 Investment	-0.244	-0.247	-0.246
	(0.145)	(0.136)	(0.137)
Constant	0.658**	0.111	0.142
	(0.216)	(0.223)	(0.215)
Observations	1,178	1,178	1,178
District Dummies	Yes	Yes	Yes
R-squared	0.518	0.514	0.497

Clustered and weighted standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3. OLS Estimation Results for Conditional Heteroskedasticity Regressions

Variables	(1) Income	(2) Income Per Capita	(3) Income Per AE
Dependent Variable: Conditional Heteroskedasticity (Standardized)			
Contract Farming	-0.238*§§	-0.223*§§	-0.225*§§
	(0.093)	(0.092)	(0.092)
Household Size	-0.005	-0.005	-0.003
	(0.020)	(0.020)	(0.018)
Dependency Ratio	-0.266	-0.189	-0.213
	(0.140)	(0.143)	(0.137)
Single	-0.024	0.041	0.048
	(0.118)	(0.120)	(0.122)
Female	0.108	0.037	0.033
	(0.115)	(0.139)	(0.140)
Migrant	0.059	0.037	0.041
	(0.071)	(0.074)	(0.074)
Age	-0.003	-0.001	-0.002
	(0.002)	(0.001)	(0.001)
Education	-0.009	-0.006	-0.005
	(0.019)	(0.019)	(0.019)
Agricultural Experience	0.005*	0.004	0.004
	(0.002)	(0.002)	(0.002)
Member of a Farm Organization	-0.042	-0.083	-0.078
	(0.057)	(0.060)	(0.061)
Taboo Days	-0.001	0.000	0.000
	(0.001)	(0.001)	(0.001)
Working Capital	0.001	-0.000	0.000
	(0.003)	(0.002)	(0.002)
Assets	0.003	0.003	0.003
	(0.003)	(0.002)	(0.002)
Landholdings	0.000	0.000*	0.000
	(0.000)	(0.000)	(0.000)
Yes to $12.50 Investment	-0.037	-0.031	-0.036
	(0.071)	(0.066)	(0.064)
Yes to $25.00 Investment	0.033	-0.015	-0.014
	(0.058)	(0.038)	(0.038)
Yes to $37.50 Investment	-0.053	-0.031	-0.029
	(0.049)	(0.051)	(0.053)
Yes to $50.00 Investment	0.021	0.056	0.054

	(0.032)	(0.058)	(0.056)
Yes to \$62.50 Investment	0.083	0.059	0.059
	(0.110)	(0.105)	(0.105)
Yes to \$75.00 Investment	0.173	0.162	0.164
	(0.166)	(0.149)	(0.151)
Constant	0.129	0.018	0.033
	(0.186)	(0.149)	(0.160)
Observations	1,178	1,178	1,178
Coefficient (Direct Effect)†	0.007	-0.012	-0.009
p-value (Direct Effect)	0.983	0.971	0.978
p-value (Joint Significance, WTP Dummies)	0.152	0.144	0.166
District Dummies	Yes	Yes	Yes
R-squared	0.071	0.067	0.065

Clustered and weighted standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1 with regular clustering
§§§ p<0.01, §§ p<0.05, § p<0.1 with Wild bootstrap
† See Acharya et al. (2016).

Table 4. OLS Estimation Results for Income Variability Regressions Exploring the Fixed Price Contract Mechanism

Variables	(1) Income	(2) Income Per Capita	(3) Income Per AE
Dependent Variable: Conditional Heteroskedasticity (Standardized)			
Proportion of Plots under Fixed Price	-0.270**§§	-0.250**§§	-0.254**§§
	(0.096)	(0.096)	(0.096)
Household Size	-0.005	-0.005	-0.003
	(0.020)	(0.021)	(0.018)
Dependency Ratio	-0.268	-0.191	-0.215
	(0.143)	(0.145)	(0.140)
Single	-0.025	0.039	0.047
	(0.116)	(0.118)	(0.120)
Female	0.105	0.035	0.030
	(0.114)	(0.138)	(0.139)
Migrant	0.053	0.031	0.035
	(0.073)	(0.075)	(0.075)
Age	-0.003	-0.001	-0.001
	(0.002)	(0.001)	(0.001)
Education	-0.009	-0.005	-0.005
	(0.019)	(0.019)	(0.019)
Agricultural Experience	0.005*	0.004	0.004
	(0.002)	(0.002)	(0.002)
Member of a Farm Organization	-0.031	-0.073	-0.068
	(0.046)	(0.049)	(0.050)
Taboo Days	-0.001	0.000	0.000
	(0.001)	(0.001)	(0.001)
Working Capital	0.001	-0.000	0.000
	(0.003)	(0.002)	(0.002)
Assets	0.003	0.003	0.003
	(0.003)	(0.002)	(0.002)
Landholdings	0.000	0.000*	0.000
	(0.000)	(0.000)	(0.000)
Yes to $12.50 Investment	-0.051	-0.044	-0.049
	(0.062)	(0.058)	(0.056)
Yes to $25.00 Investment	0.042	-0.007	-0.005
	(0.056)	(0.036)	(0.037)
Yes to $37.50 Investment	-0.054	-0.032	-0.029
	(0.049)	(0.050)	(0.051)

Yes to $50.00 Investment	0.021	0.055	0.054
	(0.030)	(0.054)	(0.052)
Yes to $62.50 Investment	0.072	0.049	0.049
	(0.111)	(0.107)	(0.107)
Yes to $75.00 Investment	0.178	0.167	0.168
	(0.169)	(0.151)	(0.153)
Constant	0.140	0.027	0.043
	(0.209)	(0.169)	(0.181)
Observations	1,178	1,178	1,178
District Dummies	Yes	Yes	Yes
R-squared	0.072	0.068	0.066

Clustered and weighted standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
§§§ p<0.01, §§ p<0.05, § p<0.1 with
Wild bootstrap

Table 5. Propensity Score Matching I: Probit Estimation Results for Participation in Contract Farming

Variables	Coefficient (Std. Err.)
Dependent Variables: = 1 if Household Participates in Contract Farming; = 0 Otherwise.	
Household Size	0.025**
	(0.012)
Dependency Ratio	-0.132
	(0.203)
Household Head Single	0.068
	(0.234)
Household Head Female	-0.449
	(0.289)
Household Head Migrant	0.066
	(0.132)
Household Head Age	-0.021**
	(0.008)
Household Head Education	-0.005
	(0.012)
Household Head Agricultural Experience	0.013
	(0.010)
Household Head Member of a Farm Organization	0.546**
	(0.213)
Household Head Taboo Days	-0.003**
	(0.001)
Household Working Capital	0.005**
	(0.003)
Household Assets	0.002
	(0.004)
Household Landholdings	0.001**
	(0.000)
Yes to $12.50 Investment	0.382
	(0.250)
Yes to $25.00 Investment	0.024
	(0.178)
Yes to $37.50 Investment	0.048
	(0.078)
Yes to $50.00 Investment	0.085
	(0.148)
Yes to $62.50 Investment	-0.213
	(0.167)
Yes to $75.00 Investment	0.401
	(0.250)

Constant	0.260
	(0.301)
p-value (Joint Significance, WTP Dummies)	0.000
Observations	1,178

Clustered and weighted standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6. Propensity Score Matching II: Treatment Effects (Three Nearest Neighbors, 0.01 Caliper)

Sample	Income	Income Per Capita	Income Per AE
Conditional Heteroskedasticity			
Unmatched Sample	-0.154***	-0.154**	-0.156**
	(0.058)	(0.058)	(0.058)
Average Treatment Effect on the Treated	-0.115**	-0.109*	-0.110*
	(0.058)	(0.057)	(0.057)
Average Treatment Effect on the Untreated	-0.220***	-0.212***	-0.214***
	(0.052)	(0.057)	(0.057)
Average Treatment Effect	-0.169***	-0.162***	-0.164***
	(0.053)	(0.053)	(0.053)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 7. Doubly Robust Weighted Regression Estimation Results for Conditional Heteroskedasticity

Sample	Income	Income Per Capita	Income Per AE
	Conditional Heteroskedasticity		
Average Treatment Effect on the Treated	-0.241§§§	-0.225§§§	-0.227§§§
	(0.078)	(0.076)	(0.076)
Average Treatment Effect on the Untreated	-0.249§§§	-0.230§§§	-0.232§§§
	(0.064)	(0.068)	(0.068)
Average Treatment Effect	-0.246§§§	-0.228§§§	-0.231§§§
	(0.068)	(0.068)	(0.068)

Standard errors in parentheses
§§§ $p<0.01$, §§ $p<0.05$, § $p<0.1$ with Wild bootstrap

Appendix

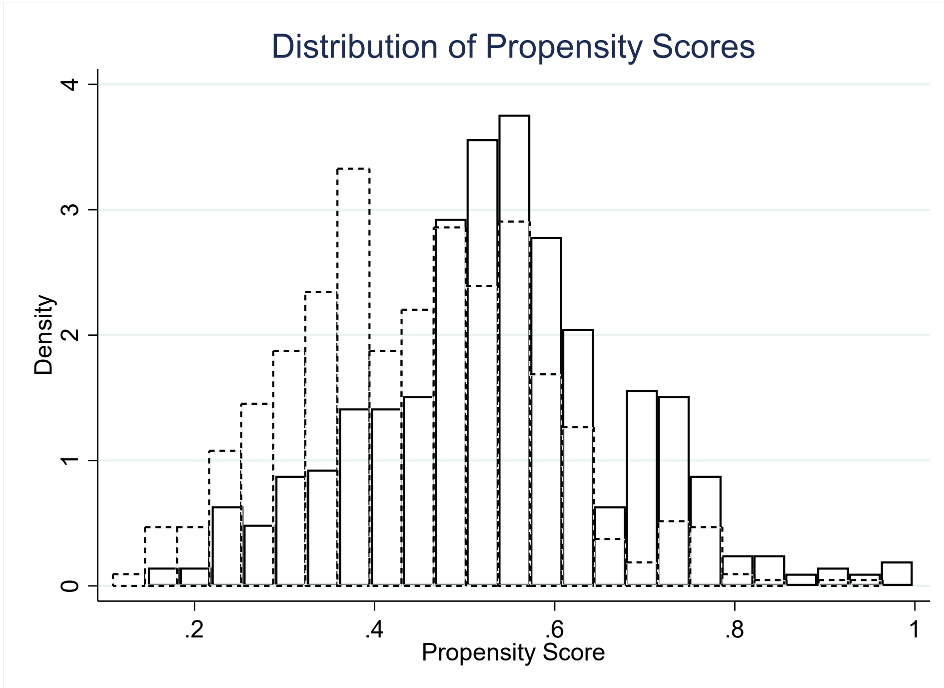

Figure A1. Distribution of Propensity Scores by Whether Households Participate in Contract Farming or Not.

Table A1. Median Regression Estimation Results for Conditional Heteroskedasticity

Variables	(1) Income	(2) Income Per Capita	(3) Income Per AE
Dependent Variable: Conditional Heteroskedasticity (Standardized)			
Contract Farming	-0.090***§§§	-0.082***§§§	-0.085***§§§
	(0.029)	(0.027)	(0.024)
Household Size	0.005	0.009	0.015***
	(0.007)	(0.006)	(0.006)
Dependency Ratio	-0.057	0.056	0.037
	(0.061)	(0.060)	(0.054)
Single	0.130	0.201***	0.219***
	(0.093)	(0.074)	(0.056)
Female	-0.070	-0.146**	-0.171**
	(0.111)	(0.068)	(0.070)
Migrant	-0.003	0.058	0.067
	(0.059)	(0.044)	(0.052)
Age	-0.001	-0.001	-0.002
	(0.002)	(0.002)	(0.001)
Education	-0.002	0.003	0.005
	(0.004)	(0.004)	(0.003)
Agricultural Experience	0.001	0.002	0.002
	(0.002)	(0.002)	(0.001)
Member of a Farm Organization	0.012	0.022	0.026
	(0.023)	(0.031)	(0.027)
Taboo Days	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Working Capital	-0.002	-0.001*	-0.002*
	(0.001)	(0.001)	(0.001)
Assets	-0.001	0.000	0.000
	(0.001)	(0.001)	(0.001)
Landholdings	0.000	0.000	0.000*
	(0.000)	(0.000)	(0.000)
Yes to $12.50 Investment	-0.032	-0.057	-0.043
	(0.045)	(0.050)	(0.045)
Yes to $25.00 Investment	0.031	0.047	0.039
	(0.046)	(0.053)	(0.042)
Yes to $37.50 Investment	0.016	0.002	-0.012
	(0.034)	(0.035)	(0.024)
Yes to $50.00 Investment	-0.002	0.021	0.059
	(0.055)	(0.044)	(0.052)

Yes to \$62.50 Investment	-0.046	-0.045	-0.070
	(0.059)	(0.049)	(0.066)
Yes to \$75.00 Investment	0.123	0.095	0.121
	(0.112)	(0.075)	(0.094)
Constant	-0.323***	-0.420***	-0.438***
	(0.085)	(0.077)	(0.072)
Observations	1,178	1,178	1,178

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
§§§ p<0.01, §§ p<0.05, § p<0.1 with
Wild bootstrap

Table A2. Western and Bloome (2009) Estimation Results for Joint Estimation of Conditional Heteroskedasticity

VARIABLES	(1) Variance Income	(2) Variance Income per capita	(3) Variance Income per AE
Contract Farming Participant	**-0.453*****	**-0.429****	**-0.435****
	(0.163)	**(0.172)**	**(0.171)**
Household Size	-0.008	0.005	0.004
	(0.019)	(0.019)	(0.019)
Dependency Ratio	-0.196	-0.181	-0.188
	(0.245)	(0.212)	(0.212)
Single	0.050	0.104	0.129
	(0.198)	(0.219)	(0.212)
Female	0.324	0.268	0.242
	(0.277)	(0.304)	(0.290)
Migrant	0.114	0.121	0.118
	(0.160)	(0.173)	(0.173)
Age	0.006	0.007	0.007
	(0.007)	(0.008)	(0.008)
Education	0.006	0.013	0.013
	(0.016)	(0.014)	(0.014)
Agricultural Experience	-0.001	-0.003	-0.003
	(0.009)	(0.010)	(0.010)
Member of Farm Organization	-0.113	-0.160	-0.154
	(0.128)	(0.139)	(0.138)
Fady Days	-0.001	0.001	0.001
	(0.003)	(0.003)	(0.003)
Working Capital	-0.002	-0.003	-0.003
	(0.002)	(0.002)	(0.002)
Assets	0.009***	0.009***	0.009***
	(0.002)	(0.002)	(0.002)
Landholdings	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)
Yes to $12.50 Investment	-0.077	-0.043	-0.049
	(0.099)	(0.091)	(0.094)
Yes to $25.00 Investment	0.067	0.001	0.003
	(0.069)	(0.057)	(0.061)
Yes to $37.50 Investment	-0.078	-0.037	-0.032
	(0.099)	(0.114)	(0.109)
Yes to $50.00 Investment	-0.005	0.036	0.039
	(0.111)	(0.135)	(0.132)
Yes to $62.50 Investment	0.444**	0.385*	0.382*
	(0.212)	(0.220)	(0.220)

Yes to $75.00 Investment	-0.168	-0.187	-0.175
	(0.246)	(0.234)	(0.240)
Constant	-0.816**	-0.976***	-0.968***
	(0.376)	(0.333)	(0.344)
District Dummies	Yes	Yes	Yes
Observations	1,178	1,178	1,178

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A3. Balance Statistics for Matched Sample with Three Nearest Neighbors and a 0.01 Caliper

	Mean				
Variable	**Treated**	**Control**	**% Bias**	**t-statistic**	**p-value**
Household Size	5.773	5.699	3.200	0.550	0.583
Dependency Ratio	0.447	0.444	1.000	0.180	0.857
Single	0.085	0.093	-2.700	-0.500	0.614
Female	0.056	0.059	-0.900	-0.170	0.865
Migrant	0.130	0.104	7.800	1.370	0.172
Age	42.555	42.714	-1.300	-0.220	0.823
Education	6.004	6.036	-1.000	-0.160	0.870
Agricultural Experience	20.102	20.145	-0.300	-0.060	0.953
Member of Farm Organization	0.269	0.279	-2.400	-0.380	0.707
Taboo Days	23.509	24.010	-1.400	-0.240	0.807
Working Capital	6.307	6.205	0.400	0.110	0.916
Assets	15.263	12.441	10.100	1.890	0.058
Landholdings	183.430	174.890	2.600	0.440	0.662
District 1	0.176	0.192	-4.100	-0.680	0.499
District 2	0.239	0.239	0.000	0.000	1.000
District 3	0.192	0.184	2.100	0.350	0.723
District 4	0.136	0.115	5.700	1.050	0.296
District 5	0.167	0.178	-2.700	-0.460	0.647
District 6	0.090	0.093	-1.000	-0.170	0.864
Yes to $12.50 Investment	0.789	0.796	-1.600	-0.290	0.770
Yes to $25.00 Investment	0.664	0.668	-0.800	-0.140	0.892
Yes to $37.50 Investment	0.486	0.492	-1.300	-0.220	0.828
Yes to $50.00 Investment	0.319	0.328	-2.000	-0.330	0.743
Yes to $62.50 Investment	0.153	0.148	1.500	0.250	0.804
Yes to $75.00 Investment	0.086	0.082	1.800	0.290	0.776