a typical model of the program participation decision are important determinants of enrollment. In the next section, I model the decision to participate in an index insurance program, first assuming that households are perfectly informed and fully trust the insurance company and then assuming that households believe the insurance company will systematically underreport shortfalls of the insured risk from the strike point.

3 A model of demand for index insurance and beliefs about the insured risk

3.1 Model structure: putting the problem in context

The model consists of farmers choosing between two activities: planting cotton on a single hectare of land, or planting a subsistence crop on the same single hectare. The subsistence crop guarantees a risk free return of *w*. Cotton is a risky technology, as output in each period is vulnerable to a covariate shock, ε^c, that is common to all households, and a households-specific shock, ε^s. In this model, I define the common shock at time *t*, ε_t^c, as the deviation of average output per hectare, or area-yields, from its mean:

$$\varepsilon_t^c = \mu - \overline{q}_t \tag{1}$$

where $\overline{q}_t$ is area-yields at time *t* and μ is the mean of $\overline{q}_t$. The price of cotton is fixed at unity, making output and revenue identical. The cotton yield of farmer *i* at time *t* is:

$$q_{it} = \mu_i - \beta_i\left(\mu - \overline{q}_t\right) - \varepsilon_t^s = \mu_i - \beta_i \varepsilon_t^c - \varepsilon_t^s \tag{2}$$

This production function is adapted from Miranda (1991). It states that in each period a farmer planting cotton receives his or her mean yield, μ_i, net of any covariate or idiosyncratic shocks, i.e., ε_t^c and ε_t^s, respectively . Note that the shocks can be harmful or beneficial, depending on the signs of ε_t^c and ε_t^s. The shocks ε_t^c and ε_t^s are assumed to be continuous, jointly independent, and

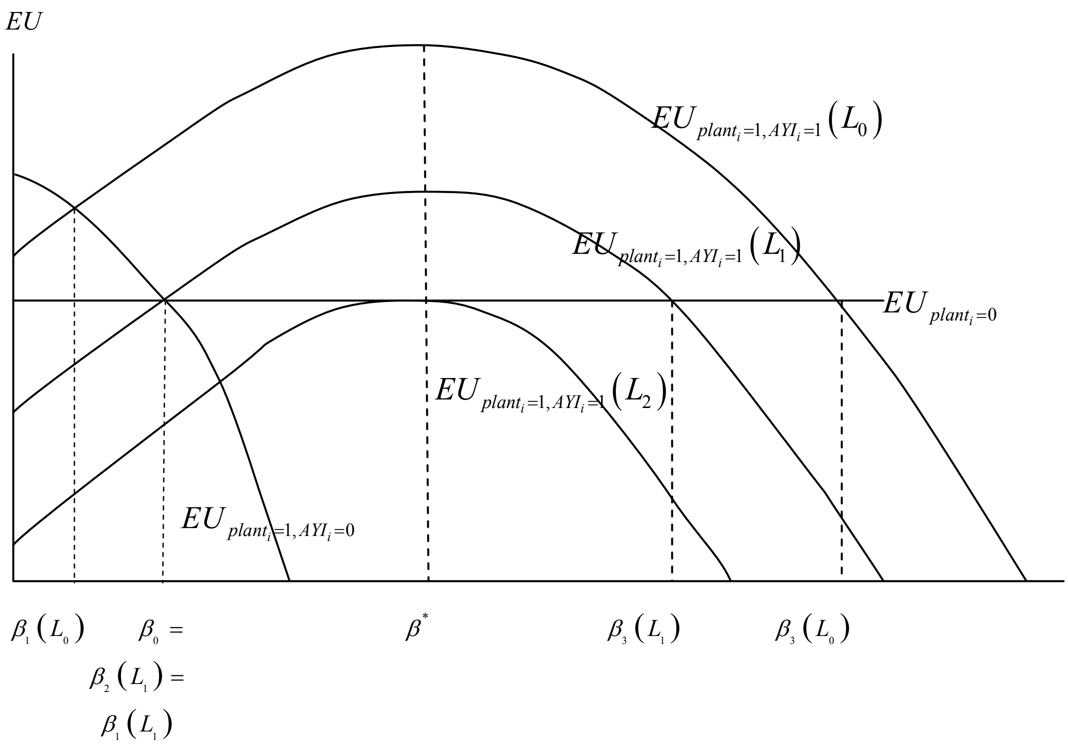

Figure 1: Expected utility of subsistence farming, cotton, and insured cotton.

The curves in Figure 1 are meant to represent the general case and are not based on any particular parameter values. Prior to the introduction of AYI, farmers choosing the subsistence crop will have expected utility located somewhere on the horizontal line segment labeled $EU_{plant_i=0}$, whereas those planting cotton have expected utility located on the curve $EU_{plant_i=1, AYI_i=1}$. Farmers with values of β_i to the left of where $EU_{plant_i=0}$ and $EU_{plant_i=1, AYI_i=1}$ intersect (i.e., to the left of β_0) plant cotton prior to the introduction of insurance.

When index insurance is introduced with a loading factor of L_0, farmers with values of β_i between β_0 and $\beta_3(L_0)$ purchase insurance and switch from the subsistence crop to planting cotton. Farmers with β_i values between $\beta_1(L_0)$ and β_0 switch from uninsured to insured cotton,

farmers) have a difficult time believing that area-yields will ever drop below the strike point. But in one sense the consequences are the same, as farmers see AYI as too expensive given the likelihood of the insurer reporting a level of $\overline{q}_t$ that would trigger a substantial payout.

Parameters values used in the simulations are given in Table 1 below:

Table 1: Parameter values for the economic model.

$\mu = 1{,}876$ kg	$\gamma = 3.5$	$\sigma_\beta^2 = 0.25$
$\sigma_c^2 = 147{,}516$ kg^2	σ_I^2=50,276 kg^2	$r = 153$ kg
$\sigma_s^2 = 147{,}516$ kg^2	$\sigma_{c,I}$=73,576 kg^2	$L = 23$ kg
$w = 1{,}565$ kg		

The values of μ and σ_c^2 were taken from a 20 year time series of cotton yields in the Pisco Valley, provided by the Ministry of Agriculture. The data were de-trended using a linear model. A normal distribution was fit to the yield data using the estimates for μ and σ_c^2 and then truncated at zero. Parameters related to the AYI contract, r, σ_I^2, $\sigma_{c,I}$, and L, were generated using the truncated normal distribution of area-yields while setting the strike point equal to zero (i.e., payouts replace 100 percent of shortfalls from μ) and loading L equal to 15 percent of the expected indemnity. In lieu of a time series of yields from agricultural households, the idiosyncratic risk parameter, σ_s^2, was set equal to the variance of the covariate shock, σ_c^2. The remaining parameters w and γ were chosen to generate a proportion of participation in cotton farming similar to what is usually observed in Pisco.

3.7 Simulation results

Simulation results depicting how perceived cheating by the insurer affects AYI demand and technology adoption are shown in Figure 4.

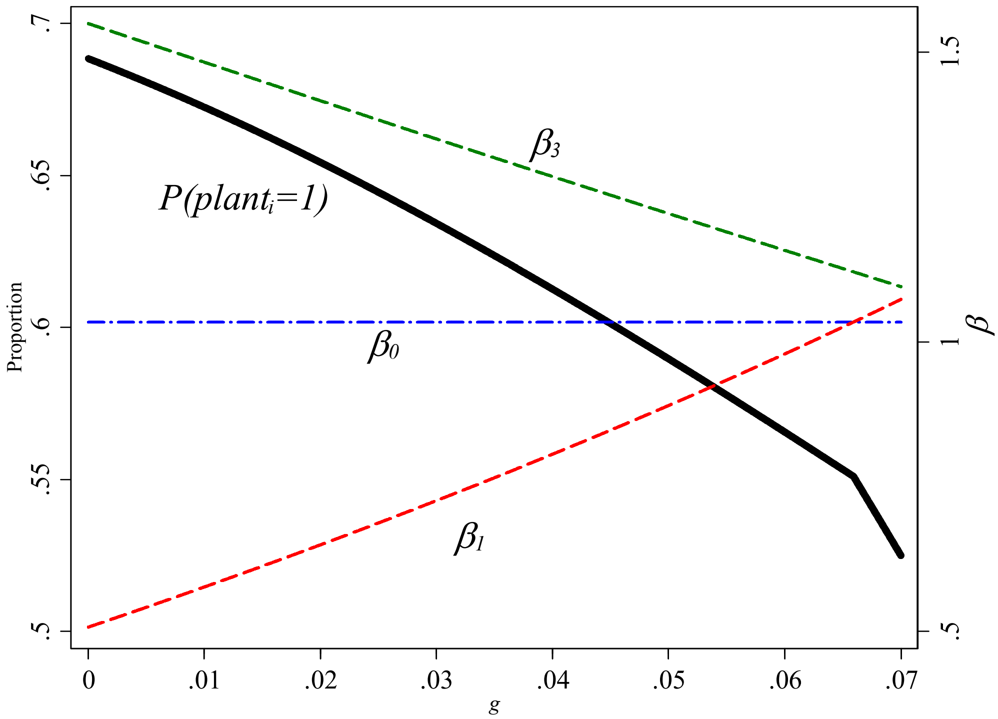

Figure 4: Perception of cheating, critical β_i values, and technology choice.

Along the horizontal axis is g, corresponding to farmer's beliefs about the percent by which insurers overstate area-yields as a proportion of μ. The middle dashed line, β_0, shows the critical value of β_i below which farmer's plant cotton prior to the introduction of AYI. The uppermost dashed line depicts β_3, gives the new upper bound on β_i for cotton planters once AYI is available and the lower dashed line is the minimum level of β_i for cotton planters, β_1. For a given value of g, farmers with values of β_i between β_0 and β_3 are induced to switch technologies from the safe activity to cotton farming, while those between β_1 and β_0 plant cotton regardless. The effect of increasing g is to narrow the range of β_i values falling between both sets of bounds, thus lowering overall demand for AYI and the proportion of farmers who are induced to plant cotton by

the incentive shifts program participation from some positive percentage of households under no extra incentive to a higher proportion, then the resulting estimate will be a Local Average Treatment Effect (LATE). The LATE will capture the average impact of the program on those that were induced to enroll by the extra incentive (Imbens and Angrist 1994).

4 Increasing demand for index insurance in the context of program evaluation

4.1 Tying the model to econometrics

Suppose we would like to estimate the average impact of AYI on technology adoption among insurance purchasers in the economy described in Section 3, and assume that the individual β_i parameters are not observed. In other words, β_i is an unobserved source of heterogeneity across farmers that will affect both the decision to participate in index insurance and whether to engage in the risky activity of cotton farming. A simple comparison of average outcomes among AYI purchasers and non-purchasers will not yield an unbiased estimate of average program impacts on participants, because as the theory from the previous sections demonstrated, the distribution of β will be systematically different across the two groups. Estimation of impacts will require an additional, observed source of variation in AYI demand that is uncorrelated with technology adoption.

4.2 A randomization of eligibility and a randomized encouragement design

One method that can yield unbiased estimates of average treatment effects is directly randomizing individuals into and out of the index insurance program. This approach may be problematic, given individual cannot usually be compelled to participate in an insurance program. Alternatively, one could randomize a variable that affects demand for index insurance without affecting the variable to be used as the outcome. Comparisons of the groups formed by this randomization can yield unbiased estimates of average treatment effects. However, the average treatment effect that is estimated may vary depending on the randomization strategy used. This is because the randomization strategy chosen will affect the composition of the comparison groups. Unless impacts are the same for all individuals or vary randomly in a way that is not correlated with the decision to buy insurance, the average outcome in each group will depend on group composition (Heckman and Vytlacil, 2007).

For example, consider a randomization of eligibility to purchase index insurance. Conceptually, this is perhaps the simplest randomization of this sort, although it may be one of the most difficult to implement in practice; for example, it might require convincing a private insurer to deny coverage to potential clients. At the individual level, randomizing eligibility would require some sort of mechanism by which randomly chosen persons are not allowed to purchase index insurance, while others are left to purchase it as they please. This is what was done by Giné and Yang (2009) in their study of demand for loans bundled with rainfall insurance in Malawi.

Denote by $z_i = 1$ random assignment of farmer i to the eligible group, i.e., those who can purchase AYI at the market price $\tau = (r + L)$, and $z_i = 0$ random assignment to the ineligible group. Suppose that a proportion $\rho_{z_i=1}$ of the eligible group buys index insurance. Using technology choice $plant_i$ as the outcome, the following average treatment effect could be estimated, using data on technology adoption and demand for insurance among the eligible and ineligible groups:

$$\frac{P(plant_i = 1 | z_i = 1) - P(plant_i = 1 | z_i = 0)}{P(AYI_i = 1 | z_i = 1) - P(AYI_i = 1 | z_i = 0)} = \frac{P(plant_i = 1 | z_i = 1) - P(plant_i = 1 | z_i = 0)}{P(AYI_i = 1 | z_i = 1)} = \frac{P_{z_i=1} - P_{z_i=0}}{\rho_{z_i=1}} \tag{21}$$

$P\left(AYI_i = 1 \mid z_i = 1\right) = P_{z_i=1}$ is the proportion of eligible farmers planting cotton, also equal to the expected value of $plant_i$ in the AYI-eligible subpopulation.

Exploiting the fact that z_i is randomly assigned and that it does not directly affect the outcome $plant_i$, it can be shown that the expression given in (21) is equivalent to:

$$\begin{aligned} &P\left(plant_{i,AYI_i=1} = 1 \mid AYI_i = 1\right) - P\left(plant_{i,AYI_i=0} = 1 \mid AYI_i = 1\right) = \\ &1 - P\left(plant_{i,AYI_i=0} = 1 \mid AYI_i = 1\right) \end{aligned} \tag{22}$$

Note that the potential outcomes $plant_{i,AYI_i=1}$ and $plant_{i,AYI_i=0}$ are used in (22), i.e., what farmer *i* would do if he were to purchase AYI, rather than the observed outcome $plant_i$. Correspondingly, $P\left(plant_{i,AYI_i=1} = 1 \mid AYI_i = 1\right)$ is the proportion of farmers who would plant cotton if they were to buy index insurance, conditional on being among the farmers that elect to buy index insurance, and $P\left(plant_{i,AYI_i=0} = 1 \mid AYI_i = 1\right)$ is the proportion of farmers in this same group who would plant cotton without index insurance.

The expression in (22) answers the following question: What is the causal effect of purchasing index insurance at the price τ on the cotton adoption rate among insurance purchasers? In the vocabulary of program evaluation, this is an example of the "Average Treatment on the Treated," or ATT. This particular average treatment effect may be of greatest relevance to policymaking, for two reasons. Firstly, it captures impacts on program participants, rather than some other sub-group. Secondly, one could also argue that the group of insurance purchasers under the randomization of eligibility ought to strongly resemble the group of farmers that will buy index insurance when it is made widely available, assuming insurance contract parameters will be the same, lending the effect estimated using the randomization of eligibility some degree of external validity. For these reasons this average treatment effect will be referred to in what follows as the "Policy Relevant Treatment Effect," or PRTE, using the terminology of Heckman and Vytlacil (2007).

Now consider a "randomized encouragement design" or allowing all households to purchase AYI but encouraging a randomly selected group to do so by offering them an extra incentive for enrollment. An example of a randomized encouragement design would be a voucher program that reduces the cost of participation in a social program.[4] Here I will use a randomized encouragement design that gives a randomly chosen subset of farmers a discount "coupon" enabling each to pay a lower price for index insurance. Denote by $c_i = 1$ assignment of farmer *i* to the encouraged group, and $c_i = 0$ if farmer *i* is not picked to receive a coupon. Suppose further that a proportion $\rho_{c_i=1}$ of the coupon group participates, while a share $\rho_{c_i=0}$ of farmers without coupons buys index insurance. This randomized encouragement yields the following estimator:

$$\frac{P\left(plant_i = 1 \mid c_i = 1\right) - P\left(plant_i = 1 \mid c_i = 0\right)}{P\left(AYI_i = 1 \mid c_i = 1\right) - P\left(AYI_i = 1 \mid c_i = 0\right)} = \frac{P_{c_i=1} - P_{c_i=0}}{\rho_{c_i=1} - \rho_{c_i=0}} \tag{23}$$

This is a "Local Average Treatment Effect" (LATE).[5] It can be shown that this expression is equivalent to:

[4] McKenzie (2009) offers some examples of encouragement designs in development.
[5] For this to represent an LATE, the encouragement must satisfy the "monotonicity" assumption (Imbens and Angrist 1994). In the case of the coupon, this assumption would require that the lower premium either encourages or has no effect on the AYI purchase decision; it cannot persuade some farmers to buy insurance and dissuade others. One could imagine some encouragement designs where satisfaction of this assumption would not be obvious. For example, exposing farmers to information about sound risk management and the role of insurance as a risk management tool might boost insurance demand, but could also directly affect technology adoption.

$$P\left(plant_{i,AYI_i=1} = 1 \mid AYI_i = 1 \leftrightarrow c_i = 1 \right) - P\left(plant_{i,AYI_i=0} = 1 \mid AYI = 1 \leftrightarrow c_i = 1 \right) \tag{24}$$

where the expression $\left(AYI_i = 1 \leftrightarrow c_i = 1 \right)$ should be read as "buys area-yield insurance if and only if given a coupon." In other words, equation (24) is the change in the cotton adoption rate due to having AYI among the group of farmers that would purchase AYI if they were to receive a coupon, but would otherwise not purchase it. This group is known as the "compliers" in the program evaluation literature.

4.3 Choosing between alternative research designs: Mean Squared Error

The fact that the LATE only captures average impacts on individuals induced to participate by the encouragement is a limitation; if individuals are heterogeneous with respect to the potential risk reduction offered by AYI, then each different possible encouragement could yield a new LATE, making interpretation of these effects difficult (Heckman and Vytlacil 2007). A corollary to this is that the effect estimated by the randomized encouragement design and that captured using the randomization of eligibility will not in general coincide. A randomized encouragement design can generate an unbiased estimate of the average impact of AYI on the group whose insurance purchase decision is determined by the encouragement, but it is a biased estimate of the PRTE when individuals heterogeneous in ways that affect program enrollment and benefits.

Bias should not be the only consideration when weighing alternative research designs, however. Changing the pool of participants via a randomized encouragement will allow households to update information with respect to the gains from participation. It seems reasonable to assume that the most effective means of allowing households to decide if they can benefit from financial market interventions is to let them experiment for themselves, or learn from the experiences of others. Randomized encouragement designs, carried out over multiple

adoption is unaffected by having insurance, then the LATE will fall. This is what we observe in Figure 6.

Larger coupons also result in a greater difference between the LATE and the PRTE. When g is at 6.7 percent of μ, the PRTE is 1; increasing the size of g has resulted in the scenario that parallels that of setting the loading factor equal to L_1 in Figure 1, i.e., the entire group of insurance purchasers if comprised of farmers induced to switch technologies from the safe crop to cotton. The LATE shrinks from 0.872 at a coupon of 1 to 0.621 when the coupon is equal to 10. The absolute bias of the LATE is initially very small, as nearly all additional farmers brought into the insurance market by the coupon only plant cotton when purchasing insurance, but increases quickly with larger coupons. In other words, increasing the value of the coupon changes the composition of the compliers to include fewer farmers whose technology adoption decision is affected by having insurance, and this composition effect grows with the size of the coupon.

5.3 Simulating the Mean Squared Error

As was stated earlier, a more complete picture of how close one can expect the estimators generated by these two competing research designs to come to the truth is given by the MSE of each. These are depicted below as a function of coupon size in Figure 7:

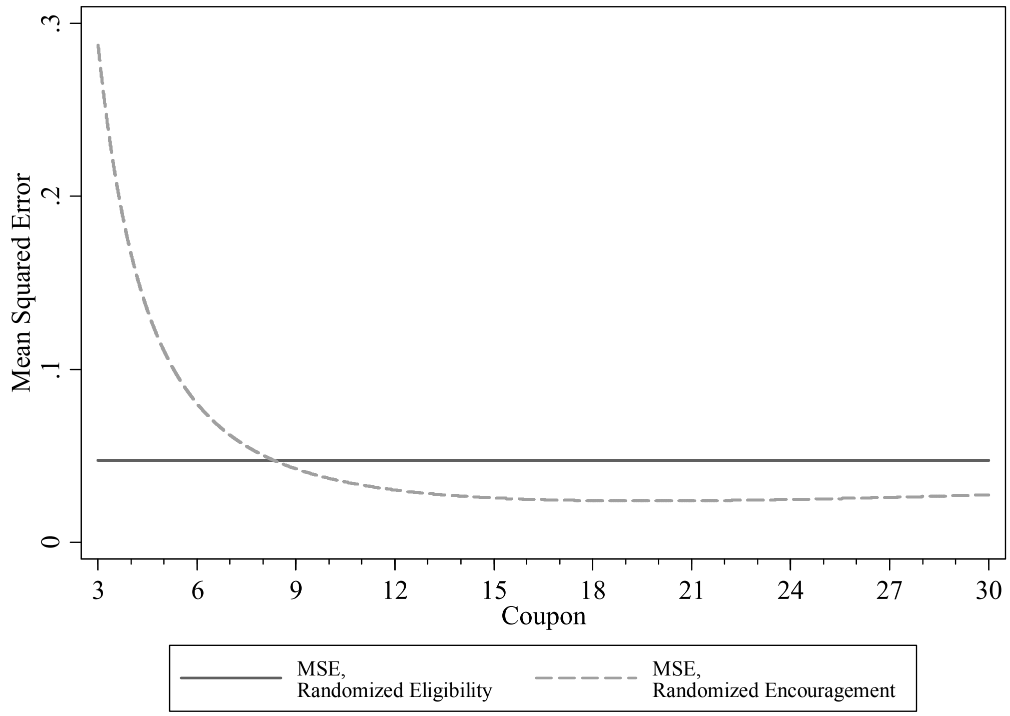

Figure 7: MSE as a function of coupon size.

Here g has been fixed at 6 percent of μ, rather than the 6.7 percent level used in Figure 6. The increase in the variance of the randomization of eligibility estimator was so sharp at $g = 0.067$ that comparing the MSE of the two estimators at $g = 0.06$ seemed more reasonable. The MSE of the estimator based on a randomization of eligibility does not change with the value of the coupon, and it is shown by the flat line located just below 0.05. At small coupon values, the impact of the encouragement on demand for AYI is very small, and as a result the MSE of the estimator for the encouragement design as depicted by the dashed line in Figure 7 is quite high. Once the coupon reaches a value of 8.35, the MSE of the encouragement design estimator drops below that of the randomization of eligibility estimator, and continues to fall until the coupon amount reaches 19.61; this is the size of the encouragement that minimizes the MSE in this case.

Beyond this coupon value, the MSE of the estimator based on the randomized encouragement begins to increase, as gains in precision diminish while bias with respect to the PRTE continues to grow.

When the MSE is used as the model choice criterion, a strong enough encouragement can make the randomized encouragement explored here preferable to a randomization of eligibility when participation is low. All of the above simulation results were generated by holding unobserved heterogeneity as represented by the spread of the β distribution, σ_{β}^{2}, constant at 0.25. A question one might ask is to what extent the degree of unobserved heterogeneity present in the population will determine the potential of a randomized encouragement design to improve upon a randomization of eligibility with respect to MSE. Under the scenario presented here, the implications of greater unobserved heterogeneity for the relative precision of estimators based on these two research designs are straightforward; a larger σ_{β}^{2} will mean that the expansion of the critical values on β_i for insurance purchasers will capture less mass of the β distribution. If participation is very low under the randomization of eligibility, then the encouragement design may still always be more precise. But it will take a stronger encouragement to generate statistically significant estimates of average treatment effects. This conclusion is, however, dependent upon the positions of the different critical values on β_i. If all were to to one side of the mean of β, for example, then a greater spread of β could increase the mass held between the critical values of β_i for insurance purchasers. Participation in index insurance would be higher under both randomization strategies.

The impact of greater unobserved heterogeneity on the difference between the estimated PRTE and the estimated LATE are less obvious. I simulate the effect of an increase in σ_{β}^{2} on the

MSE of each estimator in Figure 8 below, and then use the assumptions made about the distribution of β to gain some insight into the changes that are observed.

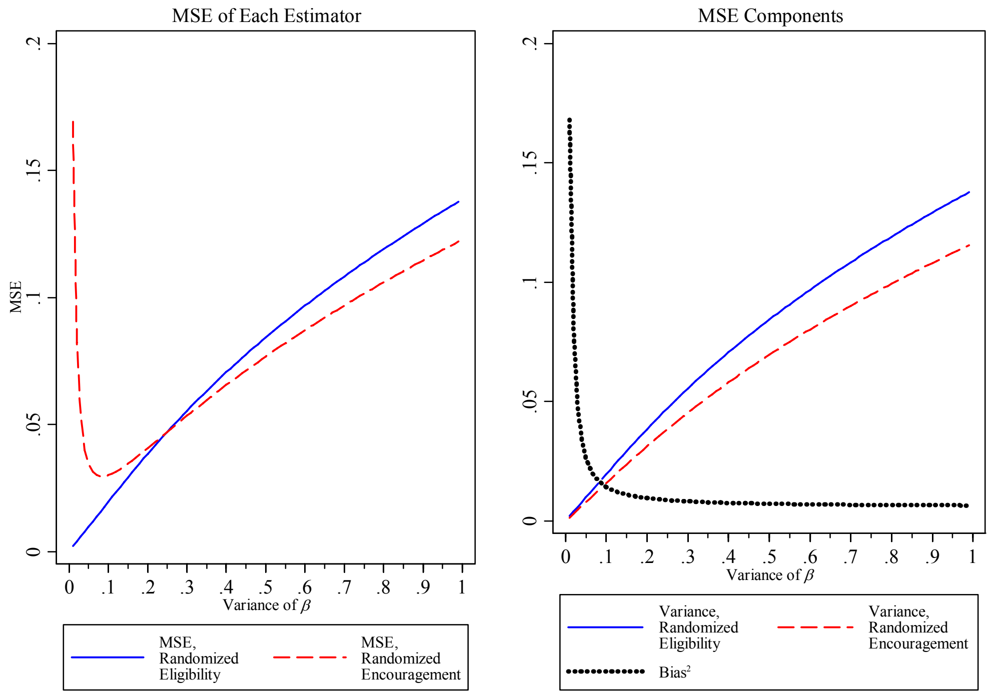

Figure 8: MSE and its components as functions of σ_β^2.

Figure 8 was generated by varying σ_β^2 while holding g fixed at 0.06, the value of the coupon at 8.35 (the value that drove the MSE of the encouragement design estimator below that of the randomization of eligibility estimator when $\sigma_\beta^2 = 0.25$), and all other parameters fixed at their levels given in . As shown in the left panel, the MSE of the estimator based on a randomization of eligibility increases steadily while that of the randomized encouragement design is minimized at $\sigma_\beta^2 = 0.09$. While the variance of each estimator grows as the spread of the β distribution increases, the change in the MSE of the randomized encouragement design estimator is offset by a reduction in the square of its bias.

The mechanics behind why the square of the bias is decreasing can be gleaned by examining the panels of Figure 9:

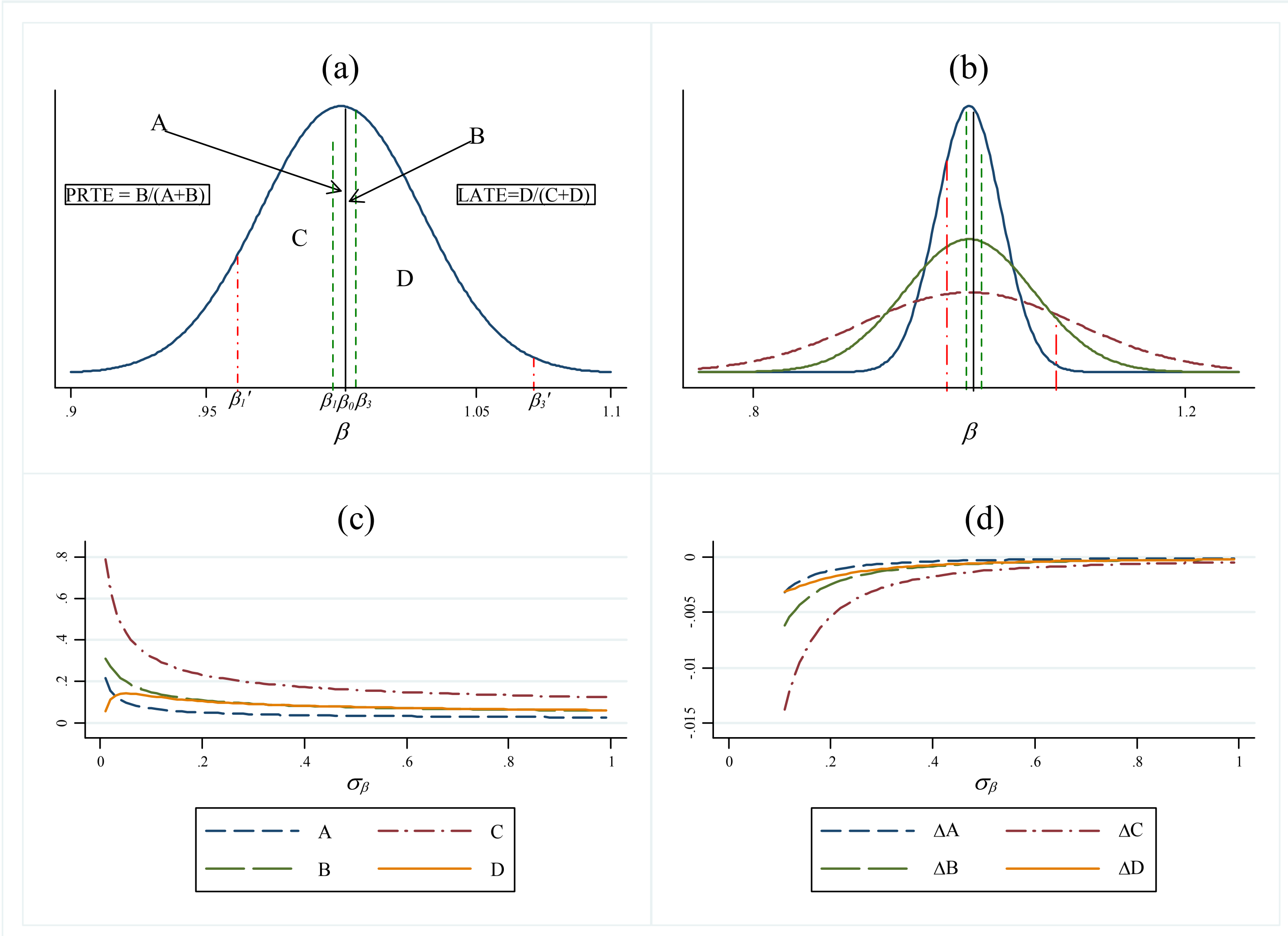

Figure 9: Change in LATE - PRTE and σ_β^2.

Panel (a) shows the areas under the curve of the density function of β that comprise the different average treatment effects. Farmers who purchase index insurance at the market price when eligible have values of β_i between β_1 and β_3, while insurance purchasers induced to plant cotton are between β_0 and β_3. The PRTE is equal to the ratio of the proportion of farmers falling into the latter group to the proportion who are part of the former group, or $\mathrm{B}/[\mathrm{A+B}]$ in Panel (a). The critical values on β_i for insurance purchasers are initially quite narrow due to the cheating parameter *g*, but are pushed out to β_1' and β_3' by the encouragement. The LATE is given by $\mathrm{D}/[\mathrm{C+D}]$, which is the proportion of compliers who are induced to plant cotton by having insurance.

Panel (b) shows the distribution of β under different values of the spread of the distribution; the density with the highest peak corresponds to the distribution in Panel (a). In general, increasing the value of σ_β^2 decreases the areas contained by A, B, C, and D, respectively. The one exception is *D* when the spread of β is very small, because of its location near the right hand tail; increasing σ_β^2 initially increases D before decreasing it. Areas holding more mass shrink more quickly, and as a result the labeled areas begin to resemble one another in size. In other words, the components of each average treatment effect are becoming more similar, and the difference in the LATE and the PRTE is diminished as a result.

This change is shown in Panel (c), which maps the change in each of the components of the LATE and PRTE against the spread of the β distribution; the masses of the different areas approach one another before their slopes level off. The face that areas holding more mass diminish more quickly is shown in Panel (d), which graphs the change in each area against σ_β.

For example, the numerator of the LATE given by area C is initially larger than all other areas, but it also has the most negative slope with respect to σ_β. As a result, its mass begins to more closely resemble that of the numerator of the PRTE, B, as σ_β increases. This is an interesting case in that it shows how a randomized encouragement design will not necessarily lead one further astray when the degree of unobserved heterogeneity in the population affecting program participation and outcomes is high. While the conclusions are specific to the model, it adds more support to the notion of randomized encouragement designs as an alternative to other strategies, or as a complement to them in the case of experiments with multiple arms.

5.4 Simulation implications

Several implications can be gleaned from the above simulation results. Firstly, larger coupons mean greater bias, and the bias-precision tradeoff involved with selecting the strength of the incentive offered by a randomized encouragement design cannot be avoided. However, bias does not necessarily translate into getting further away from the truth on average, as shown by the impact of larger coupons on the MSE in Figure 7. Randomized encouragement designs and other LATE estimators should not be dismissed out of hand because they are biased, particularly if the alternative is a research design that cannot produce much in the way of statistical precision.

Secondly, the effects of greater unobserved heterogeneity on the MSE of each estimator are context specific. Intuition might say that if a population has a great deal of underlying heterogeneity that affects both program participation and outcomes, any estimator that gets its identifying power by changing the pool of participants will be strongly biased relative to the true impact on participants. The example presented here showed that reality will often be more complex.

A1 Appendix: AYI contract parameters and comparative statics

A1.1 Demand for AYI and the parameters of the AYI contract

Recall that farmers who purchase AYI must have values of β_i lying within the following interval:

$$\max\left[\beta_1,\beta_2\right]<\beta_i<\beta_3 \tag{A.1}$$

or more explicitly:

$$\max\left[\frac{(2\mu-l)l+\gamma\sigma_I^2}{2\gamma\sigma_{I,\varepsilon_c}},\frac{\sigma_{I,\varepsilon_c}-x}{\sigma_{\varepsilon_c}^2}\right]<\beta_i<\frac{\sigma_{I,\varepsilon_c}+x}{\sigma_{\varepsilon_c}^2} \tag{A.2}$$

Where $x=\sqrt{\gamma\left(\sigma_{c,I}\right)^2+\sigma_c^2\left[\left(\mu-L\right)^2-w^2-\gamma\left(\sigma_c^2+\sigma_I^2\right)\right]}$. Assuming x is positive, which is quite reasonable given that the squared $\sigma_{c,I}$ term will be extremely large, the derivative of the upper bound of (A.2) with respect to $\sigma_{c,I}$ is:

$$\frac{1+\frac{\gamma\sigma_{c,I}}{x}}{\sigma_c^2}>0 \tag{A.3}$$

Depending on which lower bound is utilized, its derivative is either:

$$\frac{1-\frac{\gamma\sigma_{c,I}}{x}}{\sigma_c^2}<0 \tag{A.4}$$

which has an indeterminate sign, or assuming $\mu>L$:

$$\frac{-\left(\left(2\mu-L\right)L+\gamma\sigma_I^2\right)}{2\gamma\left(\sigma_{c,I}\right)^2}<0 \tag{A.5}$$

If the derivative is (A.4), the lower bound may be increasing, but at a rate slower than that of the upper bound; β_3 is always greater than β_1. Increasing the risk reduction potential of AYI via a higher value of $\sigma_{c,I}$ will weakly increase demand for insurance.

The derivative of the upper bound with respect to the variance of the indemnity, σ_I^2, is:

$$\frac{\gamma}{-2x} < 0 \tag{A.6}$$

The derivative of the lower bound with respect to σ_I^2 is then:

$$\frac{\gamma}{2x} > 0 \tag{A.7}$$

or

$$\frac{1}{2\sigma_{I,\varepsilon_c}} > 0 \tag{A.8}$$

depending on which lower bound is used. The size interval is therefore decreasing with respect to σ_I^2.

Lastly, we have the loading term L and its impact on the interval given in (A.2). The derivative of the upper bound with respect to L is:

$$\frac{-(\mu - L)}{x} < 0 \tag{A.9}$$

While the derivative of the lower bound with respect to L is:

$$\frac{(\mu - L)}{x} > 0 \tag{A.10}$$

or

$$\frac{\mu - L}{\gamma\sigma_{c,I}} > 0 \tag{A.11}$$

depending on which lower bound is used. The size of the interval is a decreasing function of the loading term, *L*.

A1.2 Perceived cheating and the AYI contract

Suppose that farmers believe that the insurer exaggerates measured area-yields in every time period by an amount $g\mu$. Given *g*, the perceived covariate shock is $\tilde{\varepsilon}_t^c = \tilde{\mu} - \tilde{\tilde{q}}_t$. The AYI contract that replaces shortfalls in area-yields from μ is perceived as:

$$\tilde{I}_t = \max\left[0, \tilde{\varepsilon}^c - g\mu\right] \tag{A.12}$$

while the true contract is:

$$I_t = \max\left[0, \varepsilon^c\right] \tag{A.13}$$

The variance of area-yields given *g* is equall to that of true area-yields. Since the common shock is symmetric, this implies that the distributions of ε^c and $\tilde{\varepsilon}^c$ are identical. When we evaluate the moments of $\tilde{\varepsilon}^c$ as they pertain to the AYI contract, we can use the distribution of ε^c as long as we account for the fact that g will leads farmers to view the strike point of the contract as higher than it actually is.

While it is proved in the main text that incorrectly perceiving the mean of area-yields leads farmers to view AYI as more costly, we cannot unequivocally say that it leads AYI to be viewed as less risk-reducing. The true change in the variance of output due to purchasing AYI is:

$$\sigma_I^2 - 2\beta_i\sigma_{c,I} \tag{A.14}$$

Replacing *I* and ε^c with $\tilde{I}$ and $\tilde{\varepsilon}^c$ yields the perceived change in variance. Consider the covariance between the indemnity and the covariate shock. The true value of this parameter is:

$$\sigma_{c,I} = E\left(\max\left[0,\varepsilon_t^c\right]\varepsilon^c\right) - E\left(\max\left[0,\varepsilon_t^c\right]\right)E\left(\varepsilon^c\right) =$$
$$E\left[\left(\varepsilon^c\right)^2 \mid \varepsilon^c > 0\right] * P\left(\varepsilon^c > 0\right) =$$
$$\left(\frac{\int_0^{\bar{\varepsilon}^c}\left(\varepsilon^c\right)^2 f\left(\varepsilon^c\right)d\varepsilon^c}{P\left(\varepsilon^c > 0\right)}\right)P\left(\varepsilon^c > 0\right) = \tag{A.15}$$
$$\int_0^{\bar{\varepsilon}^c}\left(\varepsilon^c\right)^2 f\left(\varepsilon^c\right)d\varepsilon^c$$

Under the subjective distribution of area-yields, the covariance term is:

$$\sigma_{c,\tilde{I}} = E\left(\max\left[0,\varepsilon_t^c - g\mu\right]\varepsilon^c\right) - E\left(\max\left[0,\varepsilon_{ct} - \mu g\right]\right)E\left(\varepsilon^c\right) =$$
$$\left(E\left[\left(\varepsilon^c\right)^2 \mid \varepsilon^c > g\mu\right] - g\mu E\left[\varepsilon^c \mid \varepsilon^c > g\mu\right]\right)P\left(\varepsilon^c > 0\right) =$$
$$\left(\frac{\int_{g\mu}^{\bar{\varepsilon}^c}\left(\varepsilon^c\right)^2 f\left(\varepsilon^c\right)d\varepsilon^c}{P\left(\varepsilon^c > g\mu\right)}\right)P\left(\varepsilon_c > g\mu\right) - g\left(\frac{\int_{g\mu}^{\bar{\varepsilon}^c}\varepsilon^c f\left(\varepsilon_c\right)d\varepsilon_c}{P\left(\varepsilon^c > g\mu\right)}\right)P\left(\varepsilon_c > g\mu\right) = \tag{A.16}$$
$$\int_{g\mu}^{\bar{\varepsilon}^c}\left(\varepsilon^c\right)^2 f\left(\varepsilon^c\right)d\varepsilon^c - g\mu\int_{g\mu}^{\bar{\varepsilon}^c}\varepsilon^c f\left(\varepsilon^c\right)d\varepsilon^c$$

The last line of (A.15) is larger than the last line of (A.16); i.e., the true covariance is larger than the covariance given the incorrectly perceived mean.

Ambiguity with respect to the risk reduction potential of AYI comes from the impact of g on the variance of the indemnity. The true variance of the indemnity function is equal to:

$$
\begin{aligned}
&\sigma_I^2 = E\left(I^2\right) - E\left(I\right)^2 = E\left(I^2\right) - r^2 = \\
&E\left(\max\left[0, \varepsilon_t^c\right]\max\left[0, \varepsilon_t^c\right]\right) - r^2 = \\
&E\left[\left(\varepsilon^c\right)^2 \mid \varepsilon^c > 0\right] P\left(\varepsilon^c > 0\right) - E\left[\varepsilon^c \mid \varepsilon^c > 0\right]^2 P\left(\varepsilon^c > 0\right)^2 = \\
&\left[\left(\sigma_c^2 \mid \varepsilon^c > 0\right) + E\left(\varepsilon^c \mid \varepsilon^c > 0\right)^2\right] P\left(\varepsilon^c > 0\right) - E\left(\varepsilon^c \mid \varepsilon^c > 0\right)^2 P\left(\varepsilon^c > 0\right)^2 = \\
&\left(\sigma_c^2 \mid \varepsilon^c > 0\right) P\left(\varepsilon^c > 0\right) + \left[1 - P\left(\varepsilon^c > 0\right)\right] E\left(\varepsilon^c \mid \varepsilon^c > 0\right)^2 P\left(\varepsilon^c > 0\right) = \\
&\left(\sigma_c^2 \mid \varepsilon^c > 0\right) P\left(\varepsilon^c > 0\right) + \frac{P\left(\varepsilon^c < 0\right) E\left(\varepsilon^c \mid \varepsilon^c > 0\right)^2 P\left(\varepsilon^c > 0\right)^2}{P\left(\varepsilon^c > 0\right)} = \\
&\left(\sigma_c^2 \mid \varepsilon^c > 0\right) P\left(\varepsilon^c > 0\right) + E\left(\varepsilon^c \mid \varepsilon^c > 0\right)^2 P\left(\varepsilon^c > 0\right)^2 \\
&\left(\sigma_c^2 \mid \varepsilon^c > 0\right) P\left(\varepsilon^c > 0\right) + r^2
\end{aligned}
\tag{A.17}
$$

The second to last line follows from the symmetry of ε^c, i.e., $P\left(\varepsilon^c > 0\right) = P\left(\varepsilon^c < 0\right)$.

Using similar reasoning, it can be shown that given perceived cheating by the insurer of size $g\mu$, the variance of the indemnity is:

$$
\left(\sigma_c^2 \mid \varepsilon^c > g\mu\right) P\left(\varepsilon^c > g\mu\right) + \frac{\left(\tilde{r}\right)^2 P\left(\varepsilon^c < g\mu\right)}{P\left(\varepsilon^c > g\mu\right)} \tag{A.18}
$$

Since $P\left(\varepsilon^c < g\mu\right) > P\left(\varepsilon^c > g\mu\right)$, the second term in the right hand side of (A.18) may increase the variance of the indemnity, even though $\tilde{\tau}^2 < \tau^2$. However, raising the truncation point of a distribution decreases the truncated variance, and as a result, $\left(\sigma_c^2 \mid \varepsilon^c > 0\right) P\left(\varepsilon^c > 0\right) > \left(\sigma_c^2 \mid \varepsilon^c > g\mu\right) P\left(\varepsilon^c > g\mu\right)$. The perception of cheating by the insurer unequivocally lowers the covariance between the indemnity and the covariate shock, but it may also decrease the variance of the indemnity.

A1.3 Perceived cheating by the insurer and demand for AYI

The impact of perceived cheating *g* on demand will depend on the derivatives of the bounds of the set of β_i values given in (A.2) for which purchasing AYI is optimal with respect to *g*. The derivative of the first possible lower bound of this set with respect to g is:

$$\frac{\sigma_{c,\tilde{I}}\left(2\frac{\partial h}{\partial g}\left(\mu-\left(L+h\right)\right)+\gamma\frac{\partial\sigma_{\tilde{I}}^{2}}{\partial g}\right)-\frac{\partial\sigma_{c,\tilde{I}}}{\partial g}\left(\left(2\mu-\left(L+h\right)\right)\left(L+h\right)+\gamma\sigma_{\tilde{I}}^{2}\right)}{2\gamma\left(\sigma_{c,\tilde{I}}\right)^{2}} \tag{A.19}$$

where *h* is the perceived increase the premium due to *g*. The sign of the left hand bracketed term in the numerator is ambiguous. If it is positive, then this possible lower bound is increasing with respect to *g*.

Now consider the other potential lower bound of the set, as well as the upper bound. Since $\sigma_c^2 > 0$ and $\left[\partial\sigma_{c,\tilde{I}}/\partial g\right] < 0$, the sign of derivatives of the other two bounds with respect to *g* will depend on $\partial\tilde{x}/\partial g$. This derivative is:

$$\frac{2\gamma\sigma_{c,\tilde{I}}\frac{\partial\sigma_{c,\tilde{I}}}{\partial g}-\sigma_{c}^{2}\left(2\left(\mu-\left(L+h\right)\right)\frac{\partial h}{\partial g}+\gamma\frac{\partial\sigma_{\tilde{I}}^{2}}{\partial g}\right)}{2\tilde{x}} \tag{A.20}$$

The sign of the bracketed term in the numerator is also ambiguous. Comparing the different derivatives upon which the signs of (A.19) and (A.20) depend (i.e., $\partial h/\partial g$, $\partial\sigma_{c,\tilde{I}}/\partial g$, and $\partial\sigma_{\tilde{I}}^{2}/\partial g$) yields no further information without explicitly assigning values to parameters.

References

Barrett, Christopher B., and Michael R. Carter. "The Power and Pitfalls of Experiments in Development Economics: Some Non-random Reflections." *Applied Economic Perspectives* 32, no. 4 (2010): 515-548.

Boucher, Steve, and Conner Mullally. "Case Study: Evaluating the Impact of Index Insurance on Cotton Farmers in Peru." Chap. 9 in *Designing Impact Evaluations for Agricultural Projects*, edited by Paul Winters, Lina Salazar and Alessandro Maffioli, 93-110. Inter-American Development Bank, 2010.

Bound, John, David A. Jaeger, and Regina M. Baker. "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak." *Journal of the American Statistical Association* 90, no. 430 (1995): 443-450.

Cai, Hongbin, Yuyu Chen, Hanming Fang, and Li-An Zhou. *Microinsurance, Trust and Economic Development: Evidence from a Randomized Natural Field Experiment.* Working Paper, National Bureau of Economic Research, 2009.

Carter, Michael, Peter Little, Tewodaj Mogues, and Workneh Negatu. "Poverty Traps and Natural Disasters in Ethiopia and Honduras." *World Development* 35, no. 5 (2007): 836-856.

Cole, Shawn, Xavier Giné, Jeremy Tobacman, Petia Topalova, Robert Townsend, and James Vickery. *Barriers to Household Risk Management in India.* Policy Research Working Paper, World Bank, 2010.

Deaton, Angus. *Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Development.* Working Paper, National Bureau of Economic Research, 2009.

Dercon, Stefan. "Growth and Shocks: Evidence from Ethiopia." *Journal of Development Economics* 74, no. 2 (2004): 309-329.

Duflo, Esther, Rachel Glennerster, and Michael Kremer. *Using Randomization in Development Economics Research: A Toolkit.* Technical Working Paper, National Bureau of Economic Research, 2006.

Giné, Xavier, and Dean Yang. "Insurance, Credit, and Technology Adoption: Field Experimental Evidence from Malawi." *Journal of Development Economics* 89, no. 1 (2009): 1-11.

Giné, Xavier, Robert Townsend, and James Vickery. *Patterns of Rainfall Insurance Participation in Rural India.* Staff Report, Federal Reserve Bank of New York, 2007.

Guiso, Luigi, and Tullio Jappelli. *Financial Literacy and Portfolio Diversification.* Working Paper, Centre for Studies in Economics and Finance, 2009.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. "Trusting the Stock Market." *The Journal of Finance* 63, no. 6 (2008): 2557-2600.

Hazell, Peter. "The Appropriate Role of Agricultural Insurnace in Developing Countries." *Journal of International Development* 4, no. 6 (1992): 567-581.

Heckman, James J., and Edward J. Vytlacil. *Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast Their Effects in New Environments.* Vol. 6B, chap. 71 in *Handbook of Econometrics*, edited by James J. Heckman and Edward Leamer, 4875-5143. Amsterdam: Elsevier B.V., 2007.

Hellmuth, Molly, Dan Osgood, Anne Moorhead, and Haresh Bhojwani, . *Index Insurance and Climate Risk: Prospects for Development and Disaster Management.* Vol. 2. New York: International Research Institute for Climate and Society, Columbia University, 2009.

Imbens, Guido, and Josh Angrist. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62, no. 2 (1994): 467-475.

McKenzie, David. "Impact Assessments in Finance and Private Sector Development: What Have We Learned and What Should We Learn/." *World Bank Research Observer* 31, no. 4 (2009): 209-233.

Miranda, Mario. "Area-yield Crop Insurance Reconsidered." *American Journal of Agricultural Economics* 73, no. 2 (1991): 233-242.

Rosenzweig, Mark, and Hans Binswanger. "Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments." *Economic Journal* 103, no. 416 (1993): 56-78.