The Challenges of Impact Evaluation for Agricultural Insurance Interventions

Conner Mullally
PhD Candidate
Department of Agricultural and Resource Economics
University of California, Davis
mullally@primal.ucdavis.edu

Steve Boucher
Associate Professor
Department of Agricultural and Resource Economics
University of California, Davis
boucher@primal.ucdavis.edu

Michael Carter
Professor
Department of Agricultural and Resource Economics
University of California, Davis
mrcarter@ucdavis.edu

DRAFT: Please do not cite without authors' permission

Paper prepared for the conference "On the Economics of Index Insurance," Rome., January 15-16, 2010.

Introduction

Weak or inexistent crop insurance markets are widely recognized as a significant constraint to rural development in low income countries. In response, policy-makers and researchers have proposed index insurance as a sustainable, market-based solution. Before moving forward with major insurance initiatives, however, accumulated empirical evidence on the impact of insurance is needed. To date, this evidence is scant. Giné et. al. (2008) provide evidence on the determinants of demand for the index insurance in India, home of the largest and most well-established index insurance market in the developing world. In a pilot program in Malawi, Giné and Yang (2009) find the counter-intuitive result that index insurance actually lowers credit demand and technology adoption of small farmers. Cai et. al. (2009) find a more intuitive result that a conventional (non-index) insurance for sow mortality in rural China significantly increased households' investments in sows. These research initiatives are important initial steps in the accumulation of knowledge and understanding of how index insurance works and whether or not it is likely to deliver on the promise of positive impacts on household welfare and efficiency of the rural economy. Additional, systematic research efforts are clearly needed.

This paper, which is primary methodological, seeks to facilitate future efforts. The first part takes a step back to present a general framework for causally identifying impacts of index insurance interventions. Given the difficulty of perfect compliance, in particular the practical difficulty of excluding the "control" group from purchasing insurance, we focus on an encouragement design framework. The second part of the paper draws on the lessons we have learned from a pilot project for area-yield insurance for cotton farmers that we are currently carrying out in Peru. We describe how the project came together, the specific instruments we use in the encouragement design, and some of the unexpected challenges we have faced in moving from what we believed was a well designed research project to the implementation of this project in the real world. In contrast to our expectations, demand for the insurance has been very low, jeopardizing our ability to carry out an impact evaluation. We hope that by sharing these somewhat deeper than normal insights into our research effort (including the good, the bad, and the ugly) we will contribute to more efficient research in the future on this potentially critical innovation for risk management.

1. Research Design and Econometric Evaluation of Index Insurance in Agriculture

In any evaluation of a social program, the way in which the intervention is designed is intimately connected to what sorts of conclusions the econometrician can hope to reach when conducting her analysis. For example, consider a randomized encouragement design, i.e., a factor that influences program participation but does not affect outcomes of interest is randomly varied across the population. This design will enable the analyst to estimate mean impacts for certain groups using weaker assumptions than other sorts of research designs. However, average impacts estimated from two different randomized encouragement designs applied to exactly the same population will not in general be identical, as they will describe changes in outcomes for different sub-populations. Furthermore, there is no guarantee that either of these sub-populations will be of interest to policy makers. For example, if we were interested in predicting what the average impacts of insurance would be if it were offered at market prices, a randomized insurance subsidy may not allow us to do so, as resulting estimates under relatively weak assumptions would describe behavior for individuals who would only buy insurance if offered the subsidized price.

In what follows we will use a model of participation in an agricultural insurance program and the impacts of the insurance program on an outcome of interest to motivate our examination of these questions. We consider the case of an encouragement design as described above, rather than one in which individuals are directly assigned to buying insurance or not buying. Perfect compliance, i.e., everyone in the treatment group buys insurance and no one in the control group does, seems unlikely in the case of insurance. In any case, perfect compliance is a sub-case of the model used below. The model and our examination of different types of estimable average treatment effects set the stage for a discussion of the choices that must be made with respect to research design, and how these choices will determine the set of effects that we can estimate and the segments of the population to whom these estimates are relevant. The goal is to lay out framework that will help future researchers design

1.1 A model of insurance market participation and impacts

This section lays out a basic model of demand for a particular form of index insurance, namely area yield insurance. Let y_i^U denote uninsured yields for the *i*'th farmer. Following Miranda (1991), a linear projection of individual yields on area yields results in the following equation:

$$y_i^U = \bar{y}_i + \beta_i \varepsilon_c + \varepsilon_i \tag{1}$$

Where $\bar{y}_i$ denotes the *i*'th farmer's expected yield, ε_c is the deviation of area yields from its expectation, and ε_i is, by definition, a mean zero, idiosyncratic shock uncorrelated with ε_c.

Finally, $\beta_i = \dfrac{Cov(y_i, \varepsilon_c)}{Var(\varepsilon_c)}$ measures the degree of co-movement between the individual's yield and area yield. The coefficient β_i is critical to our analysis and will be the source of unobserved heterogeneity. Since the insurance contract pays out indemnities only when area yields are low, the value of insurance to individual farmers is increasing in β_i. Basis risk, or the likelihood that the farmer suffers below average yields area yields are high (and thus no payout occurs) is decreasing in β_i.

Suppose that an area yield insurance product is made available to farmers. The area yield insurance pays farmers an indemnity equal to the shortfall of area yields from its historical mean should such a shortfall occur, and nothing otherwise. That is, $I = \max(\varepsilon_c, 0)$, where *I* is the indemnity paid to the farmer. The premium is comprised of *P*, which is the actuarially fair premium, *L* which is a loading factor charged by the insurer, and c_i, a discount on the insurance premium. This discount comes in various sizes and is offered to randomly selected farmers. Given the premium and the indemnity, income for farmer *i* should he choose to purchase insurance will be:

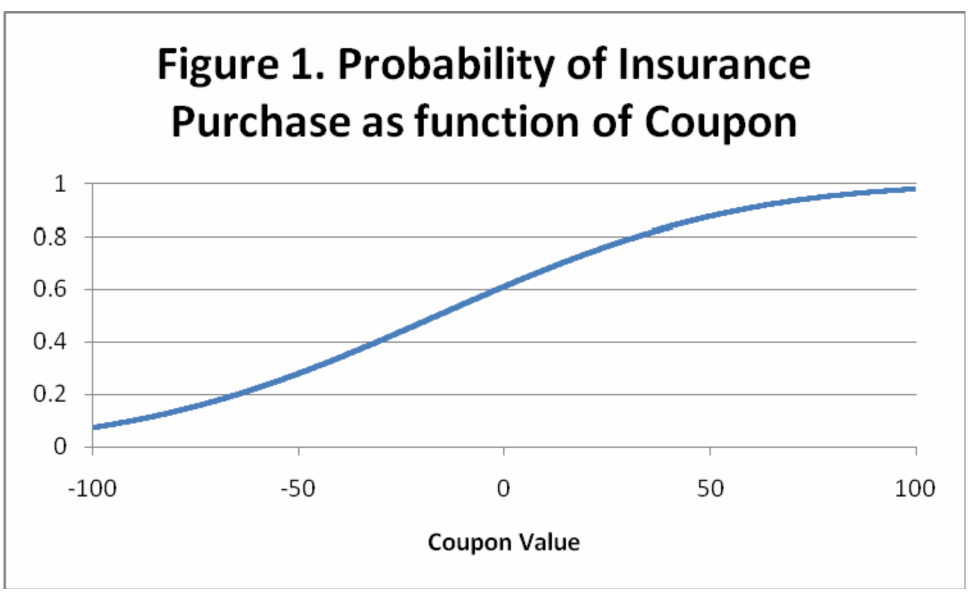
$$y_i^I = y_i + c_i - p - L + I \tag{2}$$

Farmers will purchase insurance if the expected utility gain from doing so is positive, i.e., if $EU(y_i^I) \geq EU(y_i^U)$. Assume farmers are risk averse with a mean-variance utility function: $Eu(y) = \overline{y} - \gamma Var(y)$, where γ is the coefficient of absolute risk aversion. With this utility specification, it is straightforward to show that a farmer will purchase insurance if:

$$\beta_i \geq \beta^*(c_i) = \frac{c_i - L - \gamma Var(I)}{2\gamma Cov(\varepsilon_c, I)} \tag{3}$$

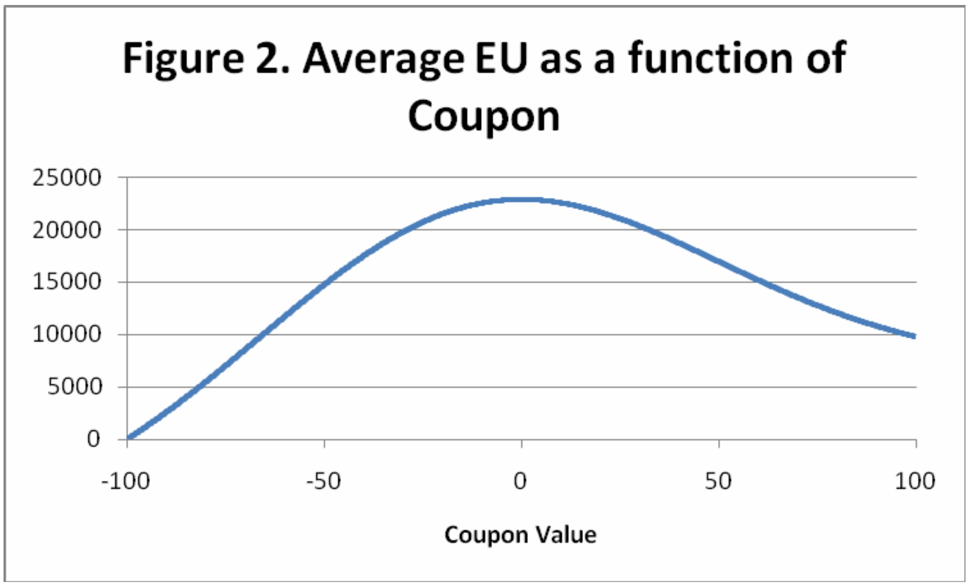
Intuitively, the greater is the value of β_i, the greater is the reduction of consumption variance resulting from insurance. Equation 3 states that a farmer will purchase insurance if their own value of β_i exceeds a threshold value which, in turn is independent of the individual's characteristics (except for the coupon value assigned to him). Variation in β_i thus generates "essential heterogeneity" across farmers; it affects both the value of insurance to an individual farmer (and thus the impact of insurance) as well as the decision to purchase insurance. Finally, note that the coupon value will be our "instrument". As seen in equation 3, by manipulating the value of coupons offered to different farmers, we can affect farmers' incentives to purchase insurance.

To drive home this point, and to foreshadow the econometric discussion, we the following two figures depict the probability of insurance purchase and the average expected utility (across farmers that purchase and don't purchase insurance) as functions of the coupon size. We set the strikepoint at 100% of the average area yield and assume loading is zero. The remainder of the parameters are based, in part, on data from cotton growers from the Pisco Valley of Peru.[1]



[1] The full set of parameter values are: …fill in later…

As implied by equation 3, Figure 1 shows that the fraction of farmers who would buy insurance increases in coupon size. The figure implies that about 60% of farmers would purchase insurance if they were offered actuarially fair insurance (with a coupon of zero). Figure 2 plots the average expected utility of farmers for different coupon values, and is just the weighted average of expected utilities with insurance for those who purchase and without insurance for those who do not purchase. We calculate expected utility with insurance *net* of the coupon value.



The intuition behind Figure 2 is as follows. Actuarially fair insurance will raise expected utility if it lowers the variance of income for the farmer. Thus, when the coupon is zero, all farmers for whom insurance lowers income variance will purchase insurance while the remainder will not. For negative coupon values, the indifferent farmer requires insurance to strictly reduce. Thus for very expensive insurance (very negative coupons), it is only those farmers with very high values of β_i who will purchase insurance. As the coupon size increase, but still remains negative, the indifferent farmer requires a smaller – but still positive -- reduction in variance to induce insurance purchase. Thus increases in coupon size to the left of zero draw in farmers who experience a variance reduction and, as a result, average expected utility increases. In contrast, further increases in coupon size beyond zero imply that the new farmers who are drawn into the insurance market actually experience an *increase* in variance as a result of insurance. With this "adverse selection" into the insurance market, average expected utility (recall that we are netting out the value of the coupon) across all farmers falls.

1.2 A model of insurance program participation and impacts

Using the above model of insurance, we now turn to empirical approaches to estimating the impact of insurance. Given that we've assumed that the only source of heterogeneity in uptake and impacts of insurance is the unobserved β_i, we proceed as if there were no other relevant covariates. The model of heterogeneous impacts outlined above suggests the use of a random coefficient framework of the type developed by Moffitt (2008). It consists of the following 2 equations:

$$Eu(y_i) = Eu(y_i^U) + [Eu(y_i^I) - Eu(y_i^U)] * d_i \tag{4}$$

$$d_i = \begin{cases} 1 & if \;\; \beta_i \geq \beta^* \\ 0 & if \;\; \beta_i < \beta^* \end{cases} \tag{5}$$

Equation 4 describes the outcome, which in our case is the farmer's expected utility which, in turn, is the expected utility from either uninsured or insured income. To streamline the presentation, let $\alpha_i = Eu(y_i^U)$ and $\Delta_i = Eu(y_i^I) - Eu(y_i^U)$, so that α_i is the outcome for individual *i* without insurance, $\alpha_i + \Delta_i$ is the outcome with insurance, and Δ_i is the change in outcome due to insurance. Of course Δ_i, the individual specific treatment effect of insurance, is unobservable because we observe either the insured or uninsured outcome for a given individual. In our impact evaluation, we are primarily interested in learning about the distribution of Δ_i.

Equation 5 describes the binary insurance purchase decision. The indicator variable d_i takes a value of 1 if the individual decides to buy insurance. The decision to purchase insurance is a function of the unobserved parameter β_i and the observed value of c_i, the value of the coupon received by the *i*'th individual

Taking conditional expectations of the model with respect to c_i, we have:

$$E[Eu(y_i) \mid c_i = c] = E\left(\alpha_i \mid c_i = c\right) + E\left(\Delta_i \mid d_i = 1, c_i = c\right) P\left(d_i = 1 \mid c_i = c\right) \tag{6}$$

$$E\left(d_i \mid c_i = c\right) = P\left(d_i = 1 \mid c_i = c\right) = P\left(\beta_i \geq \beta^*(c_i)\right) \tag{7}$$

In the left hand side of equation 6, the outer expectation is taken over the distribution of farmers. *P*() denotes probability and (6) follows because d_i is a [0,1] binary variable. We will make several additional assumptions that enable us to use c_i to identify average impacts of insurance. The first is that the assignment of the coupon, c_i , is independent of individual outcomes in the absence of insurance:

$$E(\alpha_i \mid c_i = c) = \alpha \tag{8}$$

Since the coupon is only redeemable for those who purchase insurance, randomized distribution of the coupons will ensure independence. It is important to note, however, that randomization of other instruments need not imply independence. For example, an instrument one might consider is educational sessions about insurance. If the sessions present information about yield risk in the local environment that changes farmers' priors about their own distribution of yields, the instrument could induce changes in the farmer's *uninsured* behavior.

Second, we will assume that the average impact of insurance on the outcome of interest is affected by c_i solely through the latter's affect on the fraction of the population buying insurance:

$$E(\Delta_i \mid d_i = 1, c_i = c) = g(P(d_i = 1 \mid c_i = c)) \tag{9}$$

Equation 9 is an exclusion restriction. It says that the coupon has no direct effect on farmers' utility. Instead, the average impact of insurance among purchasers with a given coupon size is solely determined by the probability that d_i=1 among that same group, i.e., the fraction of individuals assigned c_i=c that purchase insurance. Stated another way, at each coupon value, the average impact of insurance on the insured depends solely on the composition of insurance purchasers. Equation 5 tells us precisely how the composition of purchasers differs across coupon values. As the coupon value increases, $\beta^*(c)$ increases so that farmers with smaller β_i's - and thus also with lower valuation of insurance -- are induced to purchase insurance. The end result of this essential heterogeneity in impact across farmers is that the average gain from insurance among insurance purchasers is decreasing in the value of the coupon.

Next, we assume that c_i is relevant, i.e. has some predictive power with respect farmers' insurance purchase decisions:

$$Cov(c_i, d_i) \neq 0 \tag{10}$$

Assumption (10) can be tested in the sample, whereas (8) and (9) must be justified by appeals to theory. These assumptions are necessary to make c_i a valid instrument, i.e., to enable us to consistently estimate an average effect of index insurance on outcomes. Note that these assumptions are not guaranteed by randomization, as one could randomly assign an instrument c_i that also has a direct effect on the outcome value of interest with or without insurance, which would violate (8) and (9). Whether or not this is the case must be justified using theory.[2] Randomization makes for a much stronger rhetorical case, since it guarantees that the instrument is not the product of individual choice. Other instruments that are the product of individual choice seem far more likely to be correlated both with the decision to participate in a program, and unobserved heterogeneity affecting the outcome of interest.

With these assumptions in hand, we can rewrite (6) and (7) in an estimable form as:

$$Eu(y_i) = \alpha_i + g(P(d_i = 1 | c_i = c)) * P(d_i = 1 | c_i = c) + e_i \tag{11}$$

$$d_i = P(d_i = 1 | c_i = c) + u_i \tag{12}$$

The value of $Eu(y_i)$ for everyone assigned $c_i=c$ is equal to the mean outcome without insurance, α, plus the average impact of insurance among insurance purchasers in the sub-population assigned $c_i=c$, weighted by the share of this same sub-population buying insurance, plus white noise. The error terms e_i and u_i are random variables with expected values equal to 0, conditional on the probability of buying insurance.

Given the above assumptions we can estimate (11) and (12), but we need another assumption to make sure we are actually estimating an average impact of insurance on the outcome. One possibility is to assume that the effect of insurance on the outcome is the same for all subjects, i.e., a constant treatment effect. While not specific to insurance, the literature on program evaluation indicates that even after controlling for individual characteristics, this is usually not the case (Heckman and Vytlacil, 2005). There is no reason to think why insurance would be any different. The other possible assumption, which we make here, is given below:

---

[2] Examples of sensitivity analysis techniques for examining the impacts of violations of (8) and (9) on estimates are found in the section on “Assessing the Unconfoundedness Assumption” in Imbens and Woolridge (2008).

For all values of c_i such that $c_i^j \le c_i^k$, we have $\left(d_i = 1 \mid c_i = c_i^j\right) \le \left(d_i = 1 \mid c_i = c_i^k\right)$ (13)
for all *i*.

This is the "monotonicity" assumption first introduced by Imbens and Angrist (1994). The direction of the inequalities is not particularly important, as they could be reversed and the assumption would still serve its purpose. What is important is that we can order the values of the instrument c_i in such a way that moving from c_i^j to c_i^k, the sign of the impact on the decision to purchase insurance must be the same for everyone: it ether induces you to buy insurance, or it has no effect. It cannot push some people towards buying insurance, and dissuade others. If it does dissuade some others, then our estimated expected gain from buying insurance will include the expected gain among those induced to buy insurance by the instrument, minus the expected gain for those induced to not buy the insurance.[3] Thus this is an important assumption, and whether it is reasonable is not always obvious and must be considered in the context of economic theory. While we can verify that (13) holds on average within the sample by plotting the cumulative density of $P(d_i = 1)$ for the two values of c_i, we can never directly test this assumption.

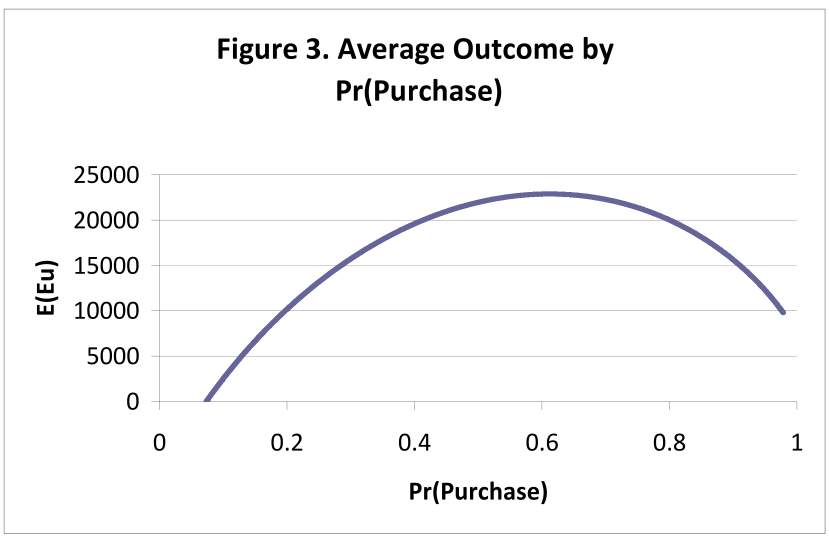
1.3 Treatment effects: General Discussion

A treatment effect for any given individual is the difference in the outcome of interest if she were to participate in the program ("treated") net of her outcome if she were not to participate ("untreated"). At any given moment, a person is either treated or untreated, and

[3] See page 434 of Angrist and Imbens (1995) for a simple explanation of the consequences for our estimated impacts of not making this assumption.

therefore we cannot observe this difference in outcomes at an individual level. Instead, the literature on program evaluation is generally concerned with estimating expected differences in these 2 outcomes for different segments of the population.

In our particular model, with data available on the instrument c_i, the outcome $Eu(y_i,)$[4] and the decision to buy insurance d_i, we could estimate the probability of buying insurance in (12), substitute this fitted probability into equation (11), and then estimate *g()* with the goal of interpreting this estimate as an average impact of buying insurance on the outcome, i.e., the "average treatment effect" for either the population as a whole, or for a specific sub-population. By varying the values of c_i, we could manipulate the probability of purchasing insurance within the population, and trace out the expected value of the outcome as a function of the probability of purchasing insurance. Figure 3 combines the information in Figures 1 and 2 to generate this function for our specific model.



In this case, the shape of the curve implies positive selection, i.e., those with the largest expected impacts of insurance are the first to choose to participate (slope is greatest at lowest probability

[4] For now we are assuming that we can directly observe farmers' expected utility (i.e., with a "utilo-meter"). We will recast the model later with a more plausible outcome variable such as farm investment or credit market participation…

of participation). As participation expands, farmers with lower expected impacts are drawn into the program, and growth in the average outcome begins to slow before finally hitting an inflection point and declining. Note that the inflection point occurs at a probability of participation of approximately 60%, which was the participation rate with actuarially fair insurance ($c = 0$). We can use this figure to illustrate the possible treatment effects that one might estimate using the model given in equations (11) and (12).

1.3.1 Marginal Treatment Effect (MTE)

If we view the expected value of the outcome variable as a welfare outcome to be optimized, then a necessary condition for a program is to be efficient is for it to equate the marginal benefits with the marginal cost of changing the expected outcome. For example, if one policy goal of an index insurance program is to increase demand for formal credit, we would want the marginal change in the probability of applying for a formal loan due to a small change in the share of the population buying index insurance to be just equal to the marginal change in the cost of the program due to changing the share of the population buying insurance. This first marginal change is the marginal treatment effect (MTE), i.e., the instantaneous change in the expected outcome in a given population due to an arbitrarily small change in the probability of program participation (e.g., buying insurance) within that same population. This concept was first introduced by Bjorklund and Moffitt (1987), and additional applications include Heckman et al. (2006), Carneiro et al. (2006), and Carneiro and Lee (2009), among others. Conceptually, the MTE is the building block of all other treatment effects, as the latter can be expressed as weighted averages of the former (Heckman and Vytlacil, 2007). In the context of our model given by (11) and (12), we can write the MTE as:

$$\frac{\partial E[Eu(y_i)]}{\partial P(d_i = 1 | c_i = c)} = g(P(d_i = 1 | c_i = c)) + \frac{\partial g(P(d_i = 1 | c_i = c))}{\partial P(d_i = 1 | c_i = c)} P(d_i = 1 | c_i = c) \tag{14}$$

Here we have continued to condition on c_i., but the MTE is defined independently of any instrument; it captures how expected outcomes shift in response to arbitrarily small changes in participation rates. In an encouragement design, these changes in participation probability come from shifts in the instrument, so we keep the conditioning notation to maintain the context of our model.

To estimate MTEs, we will need a model that generates a continuous range for the probability of purchasing insurance over some interval. Possibilities include the logit and probit models, or semi-parametric estimators can be used as more flexible alternatives; see the reviews of these methods by Ichimura and Todd (2007) and Chen (2007). A parametric model will lead to greater precision in estimation, and make it very easy to conduct welfare analysis,[5] while more flexible models will be less precise but more robust. Once predicted probabilities of buying insurance have been generated for the sample, the MTE can be estimated by picking a continuous flexible functional form for *g*(), and estimating it along with its derivative with respect to the probability of buying insurance. Standard errors can be computed by bootstrapping (14).

1.3.2: Local Average Treatment Effect (LATE)

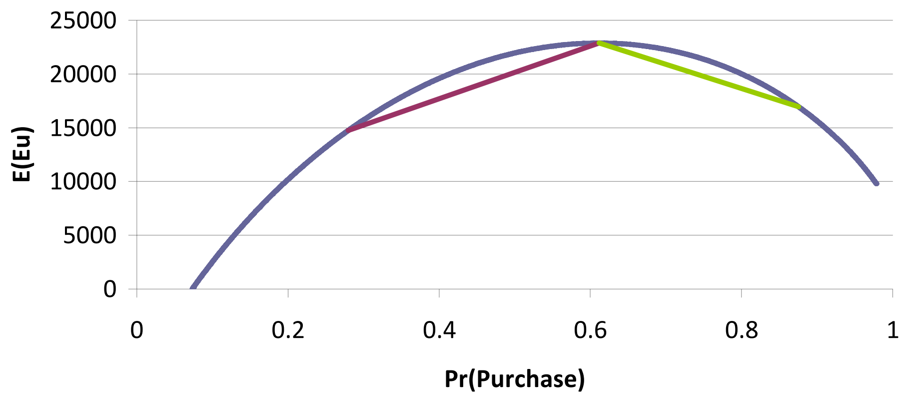
The LATE is the discrete version of the MTE, and was first introduced by Imbens and Angrist (1994). In the context of our insurance model, the LATE for any pair of instrument values (coupons) c^1 and c^2 is:

[5] For example, a logit model for the decision to buy insurance could be used to examine the average impact on consumer welfare of altering insurance contract characteristics. See pg. 55 in Train (2009).

$$\frac{E[Eu(y_i)\,|\,c_i = c^2] - E[Eu(y_i)\,|\,c_i = c^1]}{P(d_i = 1\,|\,c_i = c^2) - P(d_i = 1\,|\,c_i = c^1)} \qquad (15)$$

This is just the simple slope formula for a line segment connecting two points on the solid curve given in Figure 3. Figure 4 depicts two LATE's (associated with 3 coupon values).

Figure 4. A Depiction of LATE's

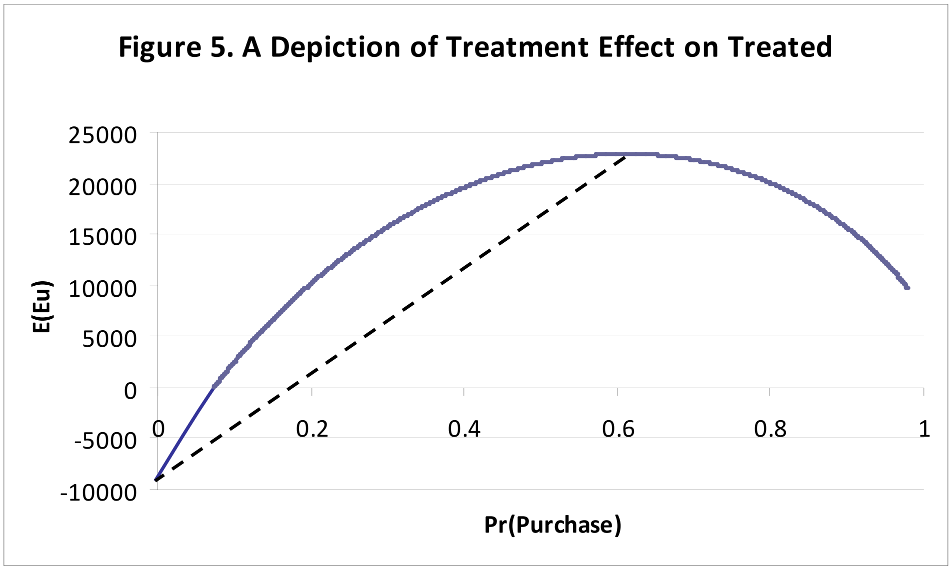


Each LATE can be estimated non-parametrically by replacing the components of (15) with sample averages. Given the monotonicity assumption, the LATE is the average impact of buying insurance on the outcome of interest for "compliers," i.e., individuals in the population who would be induced to buy insurance if assigned c^2 and would not buy insurance if assigned value c^1. In our example c_i is the amount by which the price of insurance is reduced by the coupon. Now suppose half of the sample has been randomly chosen to receive a coupon of $c^2 > 0$ (which they can only redeem if they purchase insurance), while the other half of the sample must pay the market price ($c^1 = 0$). The *compliers* are those who would buy insurance when they receive the positive coupon, but not when they pay the market price. Impacts on individuals who would always buy the insurance (the "always-takers") or would never buy insurance ("never-takers") given the values of c_i are not represented in the estimated LATE (Angrist and Imbens, 1995).

The inequalities in (13) rule out the possibility of "defiers," i.e., in this example individuals who would only buy the insurance at the market price. Observe that with a large enough support of c_i, we could approximate the shape of the function in Figure 3 using a series of LATE estimators, assuming that all pairs of values for c_i obey monotonicity (13). As long as we can reject the null hypothesis that two LATEs are equal, these estimates give us additional information with respect to the shape of the expected outcome function.

1.3.3 Treatment on the Treated (TT)

Suppose that there is some positive demand for insurance. The TT is the average effect of buying insurance on the outcome of interest for individuals who bought it. In order to identify the TT, the empirical support of $P(d_i = 1)$ must include 0 and some positive value. These two endpoints of the support can be identified non-parametrically if everyone assigned to the control group follows their assignment and does not participate in the program. For example, if we were to randomly assign farmers into two groups. Farmers assigned to the control group are not allowed access to insurance while those in the treatment group are offered actuarially fair insurance. The TT for this case is represented in Figure 5 by the dotted line segment connecting the average outcome when the probability of purchase is 0 with the average outcome that associated with the actuarially fair insurance.



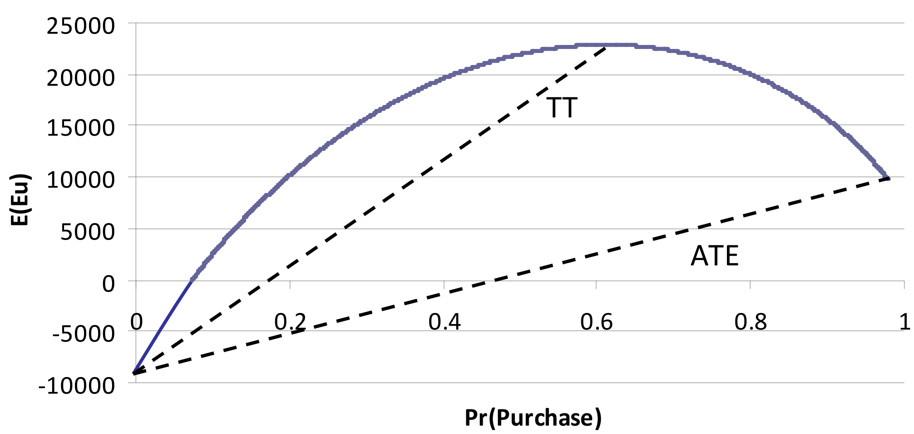
In this case, we would be able to non-parametrically estimate the TT using (15), where the second term in the denominator will now be equal to 0. Just as is the case with the MTE and the LATE, different levels of program participation will generate different estimates of TT. Also note that if we have multiple values of the instrument z_i that can be ordered in a way that satisfies monotonicity, with $P(d_i = 1)$=0 at one value of c_i , we can express the TT as a weighted average of LATEs, where the weights are proportional to the shift in the probability of participation generated by changing the value of c_i (Angrist and Imbens, 1995). This makes it possible to examine how different values of the variable c_i are affecting the overall average affect.

Estimating the TT in the case of insurance is likely to pose challenges, however, because of the practical difficulty of achieving full compliance within the control group. Specifically, insurance providers may be unwilling to deny insurance to individuals in the control group who are willing to pay for the insurance. In this case, additional assumptions must be brought to bear about the shape of the expected outcome function.

1.3.4 Average Treatment Effect (ATE)

In Figure 6, the average treatment effect (ATE) is given by the slope of the dashed line labeled ATE. In our example of an insurance program, the ATE would be the average effect of insurance on the outcome of interest if everyone in the population were to buy insurance. Identification requires that the support of $P(d_i = 1)$ include 0 and 1.

Figure 6. A Depiction of Treatment Effect on Treated



We could estimate the ATE non-parametrically under perfect compliance. Following the example above, the ATE could be estimated if everyone the coupon available to the treatment group was sufficiently high to induce all farmers to purchase insurance while the control group is again denied access to insurance (or assigned a price that is so high that nobody purchases). Observed outcomes for the two groups could be used to compute the ATE using equation (15); in this case, the denominator would be equal to 1. Again, estimating the ATE without perfect compliance is feasible, but we would have to invoke additional assumptions that permit us to extrapolate out from the empirical support of $P(d_i = 1)$ to include 0 and 1.

1.3.5 Bounds on treatment effects

Typically we are most interested in estimating either the ATE or the TT, but the support conditions needed to identify these treatment effects often fail. If we want to limit our focus to point estimation, the alternative is to estimate MTEs or LATEs, which can obviously be of interest as well. An additional option is to use the information available in the LATEs to bound the ATE and TT using techniques introduced by Manski (1990, 1997). The tightness of the bounds will depend on the assumptions we are willing to make, and the nature of the outcome variable. Let us return to our example of the randomized price discount for insurance and now let y_i now denote a general outcome variable. Suppose we randomly assign two coupons, c^1 and c^2, with $c^2 > c^1$. Furthermore, suppose that we observe that $0 < P(d_i = 1) < 1$ for $c_i = [c^1, c^2]$. The assumption of monotonicity allows us to write the ATE as:

$$\begin{aligned} ATE = P_a * (E_a(y_i \mid d_i = 1) - E_a(y_i \mid d_i = 0)) + P_n * (E_n(y_i \mid d_i = 1) - E_n(y_i \mid d_i = 0)) + \\ P_c * (E_c(y_i \mid d_i = 1) - E_c(y_i \mid d_i = 0)) \end{aligned} \tag{16}$$

and the TT as:

$$TT = P_a * (E_a(y_i \mid d_i = 1) - E_a(y_i \mid d_i = 0)) + P_c * (E_c(y_i \mid d_i = 1) - E_c(y_i \mid d_i = 0)) \tag{17}$$

Here *a* indicates always-taker, *n* indicates never-taker, and *c* indicates complier, P_a indicates the share of always-takers in the population, and $E_a(y_i \mid d_i = 1)$ is the expected outcome among always-takers with insurance. P_a is estimated using the proportion of the sample with d_i=1 and $c_i = c^1$, P_n is estimated using the proportion of the sample with d_i=0 and $c_i = c^2$, and P_c is estimated as 1-P_a-P_n. From the definition of LATE, our estimated average treatment effect using equation (15) is really an estimate of $E_c(y_i \mid d_i = 1) - E_c(y_i \mid d_i = 0)$, the average impact among the compliers.

What we do not have are estimates of the expected impact of insurance on always-takers and never-takers, $E_a(y_i \mid d_i = 1) - E_a(y_i \mid d_i = 0)$ and $E_n(y_i \mid d_i = 1) - E_n(y_i \mid d_i = 0)$, respectively. We can estimate $E_a(y_i \mid d_i = 1)$ using the expected outcome among those with $c_i = c^1$ and d_i=1, since these individuals must be always-takers. Similarly, we can estimate $E_n(y_i \mid d_i = 0)$ from the average outcome among those with $c_i = c^2$ and d_i=0, since these individuals must be never-takers. This leaves us with $E_a(y_i \mid d_i = 0)$ and $E_n(y_i \mid d_i = 1)$ as the terms that cannot be point-identified. If there are natural bounds on these expected outcomes, we can bound the ATE and TT as well. For example, suppose y_i indicates whether or not a farmer planted a hybrid seed, taking a value of 1 if this is the case and 0 if not. $E_a(y_i \mid d_i = 0)$ and $E_n(y_i \mid d_i = 1)$ must both fall within the unit interval in that case, and the upper bound on (16) can be found by setting $E_a(y_i \mid d_i = 0)$ equal to 0 and $E_n(y_i \mid d_i = 1)$ equal to 1, with the lower bound defined symmetrically, while bounds on (17) only depend on the whether we set $E_a(y_i \mid d_i = 0)$ equal to 0 or 1. If we want to further assume that farmers only buy insurance if they are better off for doing so, then $E_a(y_i \mid d_i = 0)$ is bound from above by $E_a(y_i \mid d_i = 1)$ and from below by 0, with bounds on $E_n(y_i \mid d_i = 1)$ defined symmetrically. As long as there are natural bounds on the outcome of interest, we can place bounds around the ATE and TT.

1.4 Treatment effects and research design

The point of this discussion of treatment effects is not to provide an exhaustive summary of the various types of treatment effects, the identification conditions for each, and how to bind them when these identification conditions fail. Comprehensive reviews of this sort can be found in Heckman and Vytlacil (2007) and Imbens and Woolridge (2008). Instead, the purpose is to provide key points that should be taken into consideration in the context of research design, specifically, the conditions we need to estimate valid and interpretable parameters. Consider the monotonicity assumption. If estimating average effects of insurance on the outcome is the primary goal of the study, then it is important that this assumption hold. This should be taken into consideration with respect to research design. Suppose we are choosing between two potential encouragement designs: randomly offering discount on the premium to selected farmers, or randomly inviting farmers to information sessions about the insurance product. In the first case, it is fairly clear that paying a lower price ought to (at least weakly) increase the probability of purchase, and thus the monotonicity assumption will hold. It is less clear that the monotonicity assumption will hold in the second case. One could imagine someone who might have purchased the insurance simply because her neighbors did, but by going to the information session, learned enough about the product to decide that she is better off without it. In this case, monotonicity is violated. If we were to use a combination of price variation and information sessions in our research design, we could still use the latter in our analysis. One should always check that the probability of participation is higher among information session attendees than that of non-attendees, although this does not guarantee that monotonicity holds. We would also likely want to report estimated results where we include the information session variable both in the participation and outcomes equations, i.e., condition on this variable rather than using it as an instrument, to guard against the possibility that we are not properly interpreting our estimates.

If both our instruments satisfy monotonicity, we can estimate multiple LATEs, i.e., estimated average treatment effects that capture changes in expected outcomes for a different section of the solid curve in Figure 5. A sufficient range of values for c_i would allow us to consistently estimate a discrete version of the entire curve, given a sufficient sample size at each value of the instrument. This is an important benefit of a multi-valued instrument, as it will give us an idea of how quickly the returns to a program diminish with the probability of participation, but it also points out a caveat with respect to research design. Suppose our instrument is a subsidy on the insurance premium, and the instrument c_i takes on two values: the market price for insurance, and a subsidized price that is below the actuarially fair rate. Our estimated LATEs would correspond to estimated average impacts for farmers who would buy the insurance if offered this highly subsidized price. This LATE may not be of policy interest if such highly subsidized programs are not sustainable. Picking a less generous subsidy might have allowed us to estimate LATEs for individuals who would be induced to buy the insurance by receiving a smaller discount. We want our estimated effects to tell us something about real world policies, and in the case of insurance whether or not this is possible depends strongly on the nature of the instrument used to predict demand.

This requirement must be balanced against the need to pick an instrument that will have a significant effect on demand, i.e., satisfy equation (10). This is where collecting baseline data can play a valuable role. For example, if we wanted to use random price variation as the instrument, we could include hypothetical contingent valuation questions in a baseline survey. Techniques such as those tested by Blumenschein et al. (2008) can be used to guard against the overstatement of willingness to pay typically observed in studies relying on hypothetical questions to gauge demand. This amounts to asking farmers if they would buy insurance at a particular price. If a farmer answers yes, she is asked if she is "definitely sure" or "somewhat sure" that she would buy insurance at that price. Only "definitely sure" responses are considered indicative of willingness to pay at a given price. The results of this data collection can be used to predict demand for insurance at different prices. We could also test the impact of increased information about the insurance and the interaction of better information with price incentives by explaining more about the insurance to some farmers and not to others, and doing so before and after questions about valuation.

None of this is meant to discount the value of examining the impact of a variety of factors on demand for insurance. Given the limited demand for index insurance products to this point, any information we can glean with respect to factors that influence insurance demand will be highly valuable. But if we want to combine this analysis of demand with an analysis of impacts of insurance, the nature of the instrument with respect to how it is affecting choice behavior must be taken into consideration.

1.4.1 Caveats about statistical inference in the case of an insurance evaluation

The ability to estimate LATEs or any other treatment effect hinges on two assumptions that have largely remained in the background to this point. The first is that of no general equilibrium effects, i.e., the treatment status of one member of the population does not affect the outcomes of other members of the population. This can become a concern if insurance proves to be popular. For example, we might expect farmers with insurance to use greater quantities of inputs that are “risk increasing,” i.e., inputs that increase expected profits but affect the higher moments of the distribution of profits in a way that lowers expected utility, holding expected output constant. Average treatment effects estimated from a small scale insurance project might lead to erroneous inference if a larger scale insurance program would result in higher fertilizer prices due to increased use among insurance purchasers. Another example is found in informal risk sharing networks. As insurance grows in popularity, demand for insurance among members of informal risk-sharing networks may affect transfers within the networks, which would violate the assumption of no spillovers. When possible, preliminary data collection should be done to investigate these possibilities. If it appears that this is real concern, the encouragement design must take this into account, by creating variation in demand at the level of observation which gives rise to the spill over, e.g., randomizing the instrument at the community level. However, this involves a large sacrifice with respect to statistical power, making it much tougher to precisely estimate impacts of index insurance. Therefore we would like to combine individual and more aggregate level randomization, if feasible. For example, using our example of randomized price variation, prices could be varied both at the individual level, and at whatever level is large enough to subsume the spillover. Individual price and average price at the more aggregate level would then be used as instruments, while the outcome equation would include individual demand and demand within the village, community, or whatever we have chosen as the secondary unit of randomization.

Low demand for insurance has been the norm rather than the exception; this has been our experience in Peru, and that of others in Malawi (Gine and Yang, 2008) and India (Gine et al., 2008). Thus in many cases the possibility of spillovers may not be of great concern, at least at this point. What is a concern is that our lack of understanding of what drives demand for insurance may lead us to pick instruments that have little effect on demand. If instruments have little effect on participation, estimates of average treatment effects may be biased and inconsistent (Bound et al., 1995; Staiger and Stock, 1997).[6] In this case we must use other tools to make proper inferences. One such tool is randomization inference.[7] Rather than relying on asymptotic theory as the basis of inference, randomization inference consists of constructing the

[6] Using Monte Carlo simulations, Staiger and Stock (1997) suggest as a rule of thumb that the F-statistic calculated from excluding the instruments in the first stage should be at least 10 to avoid weak instrument problems.
[7] See Rosenbaum (2002) for greater detail on randomization inference, and Greevy et al. (2004) for an application using instrumental variables.

exact distribution of a particular statistic, given a particular null hypothesis. For example, if we our hypothesis is that insurance has no affect on outcomes, then it should not matter who was encouraged to buy insurance through the instrument z_i, and the value of any statistic calculated by comparing encouraged and non-encouraged groups would be no different regardless of who was assigned to each group. For example, we could build the distribution of the difference in average outcomes between the encouraged and discouraged groups under this null hypothesis by going through every possible permutation of assignments of z_i (or as many as possible if the sample is too large), and calculating this statistic for each permutation. If a person in the group that was actually encouraged to buy insurance is assigned to the non-encouraged group in a permutation, we can use her observed outcome as what her outcome would have been in the absence of encouragement, because under the null these two outcomes are equal. We can then compare the value of this same statistic calculated using our observed assignments of z_i, and reject the null hypothesis if under the null hypothesis the probability of observing a value for the statistic as large as the one we calculated is sufficiently small.

1.4.2 Choice of instruments

Whether we use randomization inference or more traditional econometric tools, greater statistical power will always make it easier to detect significant effects. Therefore it is important to have strong instruments with respect to predicting demand for insurance. Economic theory offers some predictions as to the determinants of demand for index insurance. Much of this work is in the context of area-yield insurance, but the results are applicable to any type of index if we can decompose farm income (or yields, or consumption) into a component that is correlated with the index, and another component that is independent of the index. The work of Miranda (1991), Mahul (1999), and Vercammen (2000) derives the optimal area-yield contract under different conditions. To the extent that farmers are able to understand index insurance contracts, these papers indicate that demand should respond to changes in the premium, the trigger (i.e., the level of the index at which farmers begin to receive indemnity payments), the correlation between the index and the insured outcome (e.g., crop yield), and the shape of the indemnity function (e.g., linear versus nonlinear). This suggests that we could generate strong predictors of insurance demand by randomly varying certain contractual parameters.

However, these models assume that households understand contracts well enough to make informed decisions, and that they believe that the insurance company will keep its commitment to pay the indemnity when in index falls below the trigger. Factors that are outside the traditional framework of economic models may therefore affect demand for insurance as well. This is demonstrated by the literature on household financial decision making. Guiso et al. (2007) explore the impact of trust on participation in the stock market among Dutch and Italian households, using data from surveys of Dutch and Italian households. They find that trusting individuals are significantly more likely to buy stocks and other risk assets, and conditional on investing in a stock, invest a larger share of wealth. Cohen (2009) provides evidence that loyalty of employees to their employers leads them to over-invest retirement accounts in stock from their employers, with the average net cost of this under diversification amounting to 20% of retirement income. Guiso and Jappelli (2009) test the correlation between financial literacy, as scored using a test, and portfolio diversification, controlling for demographics. They find the two to be positively and significantly correlated. Furthermore, self-reported assessments of financial literacy and the scores on the financial literacy test are only weakly correlated. It may be that individuals not only lack the ability to make informed financial decisions, but do not recognize this shortcoming. The markets examined in these papers are not insurance markets, but they are financial markets that may not be readily understood by the average person. This all suggests that if we can manipulate the trust or loyalty felt by farmers towards insurers, or if we can affect the level of financial literacy at the individual or community level, we may create strong instruments with which to predict uptake of index insurance. How they are implemented will depend on our research goals. If an exploration of insurance demand is all we want to carry out, then we need not think quite as hard about how these factors are affecting choice as we would if we wanted to estimate impacts of insurance.

1.4.3 Choosing outcomes of interest

Models of insurance and its behavioral effects are typically static, but the reality is that impacts will vary by the time frame in which they can expect to be observed. Firstly, individuals must understand that insurance is an effective risk management tool in order to feel an incentive to make high risk, high reward investments. This is the point of the above paragraphs on "non-traditional" determinants of insurance demand, and it suggests that we might not see any impacts until general comprehension of insurance reaches a critical mass. Second, even if farmers understand how insurance works, the lag between the initial availability of insurance and when we might observe impacts will vary between outcomes. Consider the risk-rationing hypothesis put forward by Boucher et al. (2008). This hypothesis states that farmers may find themselves rationed out of formal credit markets when loan contracts require collateral, not because they do not possess adequate collateral, but because they are unwilling to bear the risk of losing this collateral. If insurance were available, these farmers could potentially redistribute a portion of this risk to the insurer, and become more willing to enter into formal credit contracts. This type of behavioral impact only requires that a person with insurance go into the office of her local lender and apply for a formal loan. However, consider the choice of cropping technology. Chambers and Quiggin (2002) make the point that the correlation between crop yield (or income) and the index, which has been shown to be the key determinant of optimal index insurance coverage, is a choice variable. Thus farmers with index insurance have an incentive to optimize their crop portfolio in a way that takes advantage of the risk reduction potential of index insurance, e.g., away from more drought resistant, low-yielding varieties and towards riskier varieties with higher expected yields. To the extent that there is a learning curve associated with new crop technologies, we might not expect to observe these sorts of behavioral effects in the short-term, or these sorts of impacts may grow in magnitude over time whereas others stay relatively constant.

These observations must be considered the context of the econometrics of dynamic treatment effects. For example, suppose we want to estimate average impacts of insurance using data from an insurance program that has existed for 2 years, and that we implemented a randomized encouragement design in the first year. Ideally, we would like to be able to estimate dynamic treatment effects, i.e., how expected outcomes are affected by the number of years with insurance. We have several options. First, we can keep our design as is, and use our single instrument to predict insurance demand in both years. The problem with this setup is that in the second year, potential outcomes (i.e., outcomes for each individual with or without insurance) may no longer be independent of the instrument; individuals in the encouraged and discouraged groups will have different probabilities of buying insurance in the initial year, which may affect their potential outcomes in subsequent years. We cannot, for example, pool observations from both years and use equation (15) to non-parametrically estimate a single LATE, whether it be static or dynamic. As Miquel (2002) proves, what we can estimate are two separate LATEs: one for those induced to buy insurance in the first period, holding the decision to buy insurance in the second period constant, and another for those induced by buy insurance in the second period, holding the decision to buy insurance in the first period constant. We can draw stronger conclusions about the nature of dynamic treatment effects, but not without making additional assumptions. Alternatively, we could re-randomize the encouragement design in the later year. This gets us back to the econometric framework described above, with the purchase of insurance in each year acting as a separate treatment, and demand for insurance in each year instrumented by receipt of encouragement in the appropriate year. This, of course, could substantially raise research costs. However, we may not much care about impacts in initial years if demand for insurance is very low, and our estimate of interest would therefore be a static impact in the later year of the program. In this case, the gains to re-randomizing the encouragement design would not be worth the associated costs.

1.5 Instruments and impacts: The empirical literature on agricultural insurance in developing countries

While the empirical literature on microinsurance is small, there are a few studies that have put some of these ideas to the test with respect to insurance demand. Cole et al. (2008) study demand for rainfall insurance in India. Their research design includes four separate treatments, randomized at the household level: price of the insurance, liquidity available to the household, sales/informational visits (with varying content), and an endorsement from a trusted agent from a local microfinance institution. Overall, 25 percent of households in the study area purchase insurance, and the majority of these purchase the lowest amount of coverage possible.

All of these treatments have significant effects, and in the expected directions. The randomized price treatment is only given to farmers who watched a video presentation about the insurance product. Cole et al. do not look at impacts of insurance, so there is no need to think about the effects of these instruments on the interpretation of estimated treatment effects. Given that the price randomization was only done among individuals who saw a video presentation, we might expect that the estimated elasticity demand only describes behavior among individuals with a certain degree of familiarity with the product, and any impacts estimated of insurance on outcomes using the premium as an instrument would also only apply to this group. This is not a criticism, it is just meant to reinforce that we must take into consideration the nature of the instruments used to predict uptake if we want to extrapolate from the results of a study to the broader population of poor farmers.

Cai et al. (2009) take a novel approach to the evaluation question in their study of a microinsurance program in China, insuring sows. The insurance product is marketed in each village by a single government marketing agent. The authors randomize the incentives faced by marketing agents with respect to selling insurance policies: some get a flat rate, some get bonuses, and the bonuses vary in size. The marketing randomization has a strong effect on demand for insurance across villages, and the authors find positive and significant impacts of buying insurance on the number of sows owned by a household. Assuming monotonicity holds, these estimates are LATEs describing average impacts on individuals induced to buy insurance by receiving a sales pitch from a more highly incentivized marketing agent. This is a village-level randomization, and the authors argue that concerns about violations of the no spillovers assumption stemming from transfers within risk-sharing networks rule out individual randomization. While there concerns are justified, individual and cluster-level randomization may complement each other within a single research design, as we described above in our discussion of treatment effects and spillovers. It is true, however, that some designs may lend themselves more to this sort of multi-level randomization than others, and it may not have been feasible in this particular case.

In their study of an index insurance product in Malawi, Gine and Yang (2008) randomize over localities whether loans made by a microfinance institution for the purchase of hybrid seeds came bundled with rainfall insurance, with the goal of testing the hypothesis that insurance should raise demand for formal credit. This is does not exactly fit in the framework we have described to this point, as the authors are measuring the impact of being offered insurance on the probability of accepting a loan, rather than measuring impacts of actually having insurance. In their case, having insurance and accepting the loan are perfectly collinear, so estimating the impact of the former on the latter would be impossible. Surprising, they find that farmers who are not offered the insurance are more likely to accept loans. They attribute this result to limited liability in credit contracts. If farmers are already implicitly insured by limited liability, then bundling insurance with the loan may amount to no more than an increase in the interest rate. This further underlines the need for greater understanding of under which circumstances we can expect the incentives provided by insurance to affect farmers in the ways predicted by theory, and what are the factors that lead to departures from more traditional models.

There are other possibilities beyond those that have already been employed. For example, one could randomly vary the point in time at which households must pay the premium. Allowing a randomly selected number of households to postpone payment of the premium could induce higher demand, and do so in a way that does not make the insurer worse off, if households have discount rates higher than the opportunity cost of funds for the insurer. In our project in Peru, we have worked with a rural microfinance institution with a large portion of its portfolio in agriculture. These loans are typically distributed in three disbursements, and currently the insurance premium is paid for out of the first loan disbursement. One could imagine designing a mechanism to randomly select which disbursement is used to pay the premium. In addition, there are contractual parameters other than the premium which might affect demand and have not been utilized as instruments. These include the index trigger, as well as the shape and slope of the indemnity function. The latter could be manipulated by designing contracts with multiple triggers; shortfalls from the first trigger would pay the farmer a certain amount per unit shortfall, while beyond the second trigger these per unit shortfall payments could increase. Availability of different contracts could be randomized over geographic units, or the intensity with which different contracts are marketed to individual clients could be randomly varied. The bottom line is that given the limited success in promoting index insurance to this point, we must be creative in choosing instruments, while thinking hard about how different instruments will affect the way in which we can interpret estimated impacts.

The point of this section has been to illustrate that choices made at every step of the research design process will affect the range of analytical possibilities available to evaluate index insurance. We believe it is important to take the following into account from the earliest stages of research design:

1. What is the population of interest, and how are they likely to respond to different randomization strategies?
2. What are the outcomes of interest?
3. Can we reasonably conclude that our design will satisfy the assumptions needed to estimate valid and interpretable parameters?

4. Will our estimated impacts describe choices made under a policy that might exist outside of our study?

Answering these questions satisfactorily is a key element of evaluating of index insurance, or for that matter, evaluating any development policy.

2. From Research Design to Implementation: Even The Best Laid Plans…

The previous discussion highlighted the importance of grounding the research design in theory. Theory guides us in identifying potential reasons for imperfections in or inexistence of insurance markets, the types of behaviors and outcomes that are likely to be impacted by strengthened insurance markets and the potential channels of impacts. Theory also helps us choose the type of impact estimate we generate and the estimation method we use, and guides us in the interpretation of the estimates we generate. Finally, theory guides our choice of instruments used to identify the causal impacts of insurance as well as to gauge the validity of the instruments themselves.

In any empirical research program, a solid theoretical foundation is necessary but not sufficient to successfully implement the research. As we have found, often rather painfully, this is particularly true in the case of index insurance. In this final section, we discuss several specific challenges we have faced in implementing an empirical research program surrounding a pilot offering of area yield insurance in Peru. Our hope is that we can contribute to the collective learning agenda about insurance research that will minimize wheel-reinventing and pitfall-stepping.

We begin by briefly describing the pilot program. Then we trace the key factors and decisions that led us to the specific program; i.e., location, institutions, contract form, etc. Finally, we discuss several unanticipated challenges that emerged in the research design.

2.1 The Pilot: Area Yield Insurance for Cotton Farmers in Pisco

In 2008, a research team from the University of Wisconsin, University of California – Davis, and the Instituto de Estudios Peruanos launched a pilot program to examine the impacts of area yield insurance on cotton farmers in the valley of Pisco, on Peru's south coast. The valley contains approximately 22,000 irrigable hectares and, as a result of Peru's agrarian reform, is dominated by small-holders who work less than 10 hectares. Cotton is the dominant crop in the valley, accounting for between 50 – 75% of planted area over the last decade. On average, approximately 60% of the 5,000 farmers in the Pisco valley plant cotton each year. The historical average yield in the valley since 1982 is 1,672 kilograms per hectare. The contract established a strikepoint of 1,412 kilograms per hectare, or 85% of average valley yield. If the average valley yield falls below the strikepoint, policy holders were paid an indemnity of $0.63 per kilogram below the strikepoint. The final premium was set at $48 per insured hectare. The actuarially fair premium was $22 per hectare. The final premium reflects loading costs (administrative costs, taxes, plus profits) of the insurance company minus a 30% premium subsidy provided by the Ministry of Agriculture.

While the research team conducted the statistical analysis underlying the insurance contract, the contract is actually offered by a triangular arrangement involving three private institutions. The insurance contract is registered is formally offered and registered with the Superintendancy of Banks and Insurance by La Positiva, one of the largest insurance companies operating in Peru. The insurance is sold in Pisco, however, by the Caja Rural Señor de Luren, a locally-based rural bank and the largest formal lender to agriculture in the region. The Caja is essentially an agent of the insurance company; collecting a small commission for each policy sold. Finally, HanoverRe provides reinsurance to La Positiva.

The agricultural year in Pisco runs from July through June. The research program will span four agricultural years. The agricultural year serves as the baseline. A random sample of 800 cotton farmers was drawn in July 2008, at the end of the 2007-2008 agricultural year. In the baseline survey, farmers were asked about their credit market participation and farm production during the 2007-2008 agricultural year. The insurance was introduced in Pisco in August 2008, and made available for the 2008-2009 agricultural year. Sample farmers were re-surveyed in August 2009 regarding their 2008-2009 campaign. Two additional years are surveying are scheduled, giving us a four year time horizon.

2.1.1 Selection of Research Site and Contract

The choice of research site was guided by several criteria. First, we required a location where sufficient data were available in order to create an index insurance contract. The data requirements, in turn, depend on the type of index considered. The first class of index contracts are area-yield insurance contracts in which the index is the observed average yield of a specific crop in a clearly defined geographic area. The second class of contracts are weather based contracts in which the index is a meteorological variable – such as rainfall, temperature, river flow, etc. – that is correlated with average yields. In either case, the first step in contract design is to estimate the probability density function (pdf) of the chosen index. The pdf can then be used to calculate the actuarially fair premium for a particular contract. Estimating the pdf requires reliable time series data; in the case of area yield index on average yields for the chosen crop within the specified area and in the case of weather based index on the particular meteorological variable being used.

In the case of Peru, we opted for an area yield insurance contract for two reasons. First, as described by Carter et. al. (2008), area yield indices tend to offer greater value (lower basis risk) to farmers than weather based indices. The intuition is that weather indices are essentially *indirectly* measured (or predicted) area yield indices. Since area yields are influenced by multiple factors, including various weather phenomenon and pest infestations, an index based on a single (or even multiple weather events) is likely to be a significantly noisy predictor of average yield, implying greater basis risk (lower value) for farmers compared to a directly measured area yield index.[8]

Second, there is a lack of reliable weather data at a sufficiently disaggregated spatial scale to construct a meaningful index.[9] In coastal Peru, virtually all agriculture is irrigated using water from rivers that originate in the Andean highlands. Most coastal valleys have some means of storing or controlling river-flows used for irrigation, making water availability endogenous and thus not viable as an index. In the highlands, micro-climates vary dramatically across very small distances, implying that basis risk would be very large for any insurance contract written on weather events measured at one of the few meteorological stations in the highlands.

After determining that we would use an area-yield instead of weather index, we next needed to determine which crop and location. Several factors guided our choice. First, we needed a crop/location with sufficient time series data to estimate the yield pdf and compute the

---

[8] A potentially important caveat is that area yield insurance may be more susceptible to moral hazard concerns than a weather-based index. Particular concern must be taken to avoid under-reporting of yields in the measurement of area yield.

[9] One exception is surface water temperatures off the Pacific coast, which are strong predictors of severe El Niño occurrences.

actuarially fair premium. We settled on a minimum time series of 15 years.[10] Next, we sought a crop-location pair with enough farmers to permit statistical inference. Given the likelihood of low insurance take-up (we initially estimated less than 25%), we sought a potential population of farmers that would be eligible for the insurance of at least 2,000. Finally, we needed a location in which a local institution would be willing to offer the insurance and collaborate with the proposed research design. These criteria reduced the potential crop-location sites to less than a handful. Ultimately, as described next, identifying the implementing institution led us to the choice of cotton in the Pisco valley.

2.1.2 Selection of Local Institutions

As discussed above, three separate institutions are involved in the provision of the insurance: the insurance company, a local financial institution, and an international reinsurance company. Our first task was to find an insurance company willing to offer the insurance. It was infeasible to offer the insurance directly (i.e., without an insurance company) for three reasons. First, the Peruvian regulatory environment prohibits non-registered insurance companies from offering any type of insurance. Second, the absence of an insurance company would require a very large research budget as it would have to cover the maximum potential indemnity payouts. Third, and most importantly, we sought to maximize external validity of the research project. One way to do this is to have the insurance offered under market conditions and by the same types of market institutions that would offer it if the insurance were to be scaled up.

Finding an insurance company partner in Peru turned out to be somewhat of a challenge. The main reason is that the formal insurance market in Peru is highly concentrated in the major urban areas and on traditional policies of home, life and auto. It is only in very recent years that

[10] 15 years was a relatively back of the envelope calculation. We were somewhat surprised in our conversations with several international reinsurance companies who stated that they were willing to underwrite contracts with as few as 10 years of data.

insurance providers have become interested in rural areas. The insurance market for agriculture is particularly thin. There are likely many reasons for this, including the high degree of informality in the agricultural economy, the high transactions costs and ensuing large information asymmetries in rural areas, the predominance of small-scale and, in many cases, non-commercial farmers, and the historically high degree of political risk in the form of state incursions in the private markets. A long period (15 years) of relative stability and growth as well as a marked deepening of formal rural financial markets has raised the interest of several insurance providers in this potentially large and untapped market.[11] In order to identify potential insurance partners, we worked through APESEG, the Peruvian Association of Insurance Companies. APESEG organized several meetings with representative of the various insurance companies to whom we introduced the concept of area yield insurance and proposed a research partnership. Several companies immediately dismissed the idea because they were not interested in crop insurance. After nearly 18 months of conversations (and to our great relief), one company, La Positiva, decided to participate in the pilot program.[12]

Two factors contributed to the need for an additional, local institution to deliver the insurance. First, consistent with the historical lack of participation of insurance companies in agriculture described above, La Positiva had no experience offering crop insurance. Second, as discussed in the first part of the paper, insurance requires a significant degree of trust by the farmer in the insurance provider. La Positiva thus needed a local partner that both understood cotton farming and had a solid reputation among cotton farmers. The Caja Rural Señor de Luren

[11] An additional incentive was provided by President Alan Garcia who, in his electoral campaign in 2006 proclaimed that agricultural insurance would be a policy priority of his administration. In 2008, the government approved the creation of a $10 million insurance fund that would be used to support insurance initiatives from the private sector. Until now, the insurance contract designed for this research project is the only initiative to emerge.
[12] As is likely typical in this type of research project, the key was to find the individual manager who was willing and able to think “outside the box” and take the leap of faith of working with academic researchers.

met both of these requirements. The Caja has offered financial services in Pisco for over 20 years and is by far the largest provider of credit to cotton farmers in the valley. After several meetings with the board of directors, the Caja leadership agreed to participate in the pilot program.

The final institutional actor is the international reinsurance company, HanoverRe. As La Positiva had worked extensively in the past with HanoverRe, they opted to use their reinsurance services for this policy as well. The reinsurer is critical in this type of pilot project in which risk is not well diversified for the local insurance company. La Positiva chose to offload approximately 80% of the risk associated with the contract in the first two years.

2.1.3 The Encouragement Design

In negotiations with the Caja and La Positiva, we decided to use an encouragement design in order to identify the impacts of the insurance. Understandably, the Caja was unwilling to deny access to the insurance to existing or new clients because they were randomly assigned to a "control" group. The encouragement design utilizes two separate instruments: coupons and educational sessions. The coupons allow farmers to receive a discount on the insurance premium and thus are equivalent to a randomization over price. Three different coupon values were distributed: 15 S/., 35 S/., 65 S/., and 90 S/. per insured hectare. The 35 S/. coupon lowered the effective price of the insurance to just below the actuarially fair rate. We randomly selected 543 cotton farmers to receive the coupons, with one fourth receiving each value.[13]

The educational sessions, which are described in more detail below, had two levels of randomization. First, we randomly chose 16 of the 40 irrigation sub-sectors in which the valley is divided. In each of the sub-sectors, we carried out a single educational session. Within each sub-sector, we delivered invitations to the session to 60 randomly selected cotton farmers. Only

[13] The cost of the redeemed coupons is assumed by the research team.

invited individuals were allowed to participate. Approximately 75% of invitation recipients attended the sessions.

2.2 Unanticipated Research Challenges

As researchers, we were quite disappointed when we learned that insurance take-up, or purchase, was minimal by the closing sales date of the first year. Only 52 policies were sold overall, of which only 12 were within our sample. Despite our best laid research plans, we had essentially no means to estimate the impact of insurance. As with most empirical research, we were confronted with several unforeseen challenges that, we believe, were partially responsible for the low take-up rates. We discuss three of the primary challenges we faced and, when possible, how we have adjusted our strategy accordingly.

2.2.1 Educational Sessions: How much learning really occurred and are they a valid instrument?

The educational sessions were designed to meet two goals. First, they are designed to teach farmers about index insurance so that they can take an informed purchase decision. Insurance is complex because it is a state contingent, inter-temporal contract. Premiums are paid up-front, however indemnities are only paid to the farmer if certain conditions are met. Index insurance adds implies an additional complexity, namely that the conditions under which indemnities are paid are not directly related to the farmers' agricultural outcome, but instead to an external index. Conveying the basics of insurance contracts to the "never-before-insured" is a non-trivial matter. The strategy we utilized within the educational sessions drew on experimental economics.[14] We created a simple game designed to simulate stochastic farming outcomes with and without insurance. Farmers' net income in each round of the game depends on their project choice,

[14] For a detailed description of the game protocols and results, see Galarza (2009).

whether or not they purchase insurance and two types of shocks: a weather shock that affects average yields throughout the valley and an idiosyncratic shock specific to the farmer. The insurance made a payment to the farmer if the weather shock was bad (simulated by drawing a black poker chip from a bag). To introduce the concept of basis risk, the insurance payment was independent of the idiosyncratic shock (simulated by drawing different colored ping pong balls from a bag). To promote experimenting and learning about insurance, farmers played for 12 "seasons" or rounds. The amount of real money farmers earned depended on their performance in the games, with average earnings amounting to 2 days of the local agricultural wage.

While we believe that the use of games has high potential for teaching farmers about index insurance, it appears that the games were less effective than we initially hoped at conveying the basics of the insurance contracts. At the end of the game sessions, a brief survey was administered to farmers. Several questions were asked to judge farmers' understanding of the index insurance contract presented in the game. One question asked if the indemnity payout depended on the farmer's draw from the idiosyncratic shock bag. Just over a quarter of the farmers incorrectly said that it did. A more challenging issue emerged in the open question and answer session after the games were completed. Many farmers whose own average yields were higher than the valley average were quite skeptical about buying the insurance because their own yields were very unlikely to fall below the strikepoint. These comments again suggest that even after the educational session, these farmers did not fully understand how index insurance functions. Additional work is needed in refining the sessions, in particular to convey the notion that the value of the insurance depends on the degree of co-movement between an individual farmer's yields and the index (in this case valley average yield).[15]

[15] An entirely different, although perhaps no less important, challenge is the notion of "average". Many farmers seem to equate "average yield" (*rendimiento promedio*) with the parcel's potential yield or the yield they would

The games also were intended to serve as a second instrument within the encouragement design, with the notion that participation in the games would affect (increase) the probability of farmers purchasing insurance while having no direct impact on the farmers' behavior with or without insurance. There are two reasons that the games might have violated the assumptions required of an instrument. First, the sessions may have had a direct impact on behavior. For example, the in-depth discussions about both covariate and idiosyncratic risk in the valley may have affected farmers' perceptions of the risk they face (i.e., modified their subject yield pdf's). If so, then even if they don't purchase the insurance, they may adjust the way they grow cotton (or the amount of cotton grown). Similarly, if they do purchase insurance, their input decisions may be different compared to if they had purchased insurance but not attended the education session.

The second assumption that could be violated is monotonicity. Recall that monotonicity requires that the education session should either weakly increase the probability of insurance purchase for all participants or weakly decrease it for all participants. Given the structure of the sessions, however, it is possible that the impact of the sessions works in opposite directions for some individuals. For example, a farmer who had a good experience in the sessions, winning a relatively large amount of money, may be encouraged to purchase the insurance while a farmer who fared relatively poorly may develop a negative association with the insurance and be less inclined to purchase it (compared to if he had not attended the educational session). One could imagine other potential interactions of games with farmer characteristics – such as education -- that would lead to opposite directions in the impact of the educational sessions on the probability

expect to get in a good year. Similarly, farmers seem to discount exceptionally bad years from their mental calculation of average yields. This factor tend to make farmers' perceptions of average yields (both their own and in the valley) quite a bit higher than the statistical mean. Equivalently, farmers' subjective pdf's of average valley yields may be shifted significantly to the right relative to the pdf generated from official statistics. If this is the case, then farmers would underestimate the value of the insurance and be less likely to buy it.

of purchasing insurance. As discussed above, if we had strong reason to suspect that monotonicity does not hold, we would be unable to use the sessions to increase the density of the support of the probability of participation, however we would use participation in the sessions (and perhaps individual outcomes within the sessions) as conditioning variables in the analysis.

2.2.2 Understanding Institutions' Incentives

A second problem that emerged rather unexpectedly was a resistance to the project from the manager of the local branch of the Caja Rural. In hindsight, our mistake was not developing a sufficient understanding of the incentives in the management system in the Caja. As mentioned above, the research team negotiated a participation agreement with the board of directors of the Caja. The board of directors, who meet in the Departmental capital in the city of Ica, then passed the decision down to the manager of the Caja branch in Pisco. For a number of reasons, this manager was not enthusiastic with the project. First, the manager likely resented the process; the order to implement an experimental pilot program came down from above without any input from the branch itself. Not only was the order to participate very vertical, but it also implied some costs in terms of time and training for the loan officers who would be the face of the insurance product. Second, the board of directors ordered an interest rate cut on loans for farmers who also purchased the insurance. As we later learned, the manager resented this because he felt it reduced the branch's earnings. Although in the long run, insurance would likely reduce default rates and thus offset the interest loss, the manager was – understandably -- concerned with the short run earnings position of his branch. In addition, communication between the insurance company and the Caja was less than optimal. For example, there was confusion about who would lead the marketing campaign for the insurance product. By the time the confusion was cleared up, the insurance sales season had already begun.

We have attempted to address these problems by increasing communications flow between all parties and, in particular, by attempting to create incentives for the manager to fully get behind the project. In part this has taken the form of including the manager and the loan officers in discussions about how to improve the product, making it more valuable to their clients. We have also provided monetary incentives for loan officers and the manager based on the number of policies sold.

2.2.3 Macro and Political Economy Concerns

Two final problems emerged after we initiated the research project. First, the García administration investigated plans for alternative, national crop insurance programs. Although no concrete program has yet to emerge, many farm groups were expecting the government to announce a comprehensive public insurance program at highly subsidized prices. As a result, some farmers may have been reluctant to purchase the relatively expensive area yield insurance contract offered in our pilot study.

Second, two adverse macro economic shocks impacted cotton production. First, the 2008 oil price shock dramatically increased key input prices. Cotton in Pisco is heavily dependent on chemical fertilizers (average spending per hectare on fertilizers in Pisco is near $200). Second, the Peru Trade Promotion agreement went into effect in February, 2009. The reduction in trade protection strongly impacted cotton, reducing prices by 20 – 30%. As a result of these two shocks, many farmers have reduced their area planted to cotton or switched out of cotton production altogether. Since it is tied to area yield insurance for cotton growers, the project could be jeopardized by a significant reduction in area planted or the number of cotton farmers.

3. CONCLUSIONS

…Coming soon!...

References

Angrist, J., Imbens, G., 1995. "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity." *Journal of the American Statistical Association*, 90(430), 431-442.

Bjorklund, A., Moffitt, R. 1987. "The Estimation of Wage Gains and Welfare Gains in Self-Selection Models." *The Review of Economics and Statistics*, 69(1), 42-49.

Blumenschein, K., Blomquist, G. Johannesson, M., Horn, N., Freeman, P. 2008. "Eliciting Willingness to Pay without Bias: Evidence from a Field Experiment." *The Economic Journal*, 118(525) 114-137.

Boucher, S., Guirkinger, C., Carter, M. 2008. "Risk Rationing and Wealth Effects in Credit Markets: Theory and Implications for Agricultural Development." *American Journal of Agricultural Economics*, 90(2), 409-423.

Bound, J., Jaeger, D., Baker, R. 1995. "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(), 442-450.

Cai, H., Chen, Y., Fang, H., Zhoi, L. 2009. "Microinsurance, Trust, and Economic Development: Evidence from a Randomized Natural Field Experiment." NBER Working Paper 15396. http://www.nber.org/papers/w15396. Accessed October 15, 2009.

Carneiro, P., Heckman, J., and Vytlacil, E. 2006. "Estimating Marginal and Average Returns to Education." Unpublished working paper. http://www.ucl.ac.uk/~uctppca/school_all_2006-10-30b_mms.pdf. Accessed September 15, 2009.

Carneiro, P., Lee, S. 2009. "Estimating Distributions of Potential Outcomes Using Local Instrumental Variables with an Application to Changes in College Enrollment and Wage Inequality." *Journal of Econometrics*, 149(), 191-208.

Chambers, R., Quiggin, J. 2002. "Optimal Producer Behavior in the Presence of Area-yield Crop Insurance." *American Journal of Agricultural Economics*, 84(2), 320-334.

Chen, S. 2007. "Large Sample Sieve Estimation of Semi-nonparametric Models." In Heckman, J., Leamer, E. (Eds.), Handbook of Econometrics, vol. 6B, 5549-5623. Elsevier, Amsterdam.

Cohen, L. 2009. "Loyalty-based Portfolio Choice." *Review of Financial Studies*, 22(3), 1213-1245.

Giné,. X., Townsend, R., Vickery, J. 2008. "Patterns of Rainfall Participation in Rural India." *The World Bank Economic Review*, 22(3), 539-566.

Gine, X., Yang, D. 2008. “Insurance Credit, and Technology Adoption: Field Experimental Evidence from Malawi.” *Journal of Development Economics*, 89(1), 1-11.

Greevy, R., Silber, J., Cnaan, A., Rosenbaum, P. 2004. “Randomization Inference with Imperfect Compliance in the ACE-Inhibitor After Anthracycline Randomized Trial.” *Journal of the American Statistical Association*, 99(465), 7-15.

Guiso, L., Jappelli, T. 2009. “Financial Literacy and Portfolio Diversification.” Centre for Studies in Economics and Finance, Working paper 212. http://www.csef.it/WP/wp212.pdf. Accessed October 2, 2009.

Guiso, L., Sapienza, P., Zingales, L. 2007. “Trusting the Stock Market.” Forthcoming in *The Journal of Finance*. http://www.bus.wisc.edu/finance/Meyferth/Fedenia/Articles/Guiso%20Sapienza%20Zingales%202007.pdf. Accessed October 1, 2009.

Heckman, J., Urzua, U., Vytlacil, E. 2006. “Understanding Instrumental Variables in Models with Essential Heterogeneity.” *The Review of Economics and Statistics*, 88(3), 389-432.

Heckman, J., Vytlacil, E. 2007. “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.” In Heckman, J., Leamer, E. (Eds.), Handbook of Econometrics, vol. 6B, 4875-5143. Elsevier, Amsterdam.

Ichimura, H., Todd, P. 2007. “Implementing Nonparametric and Semiparametric Estimators.” In Heckman, J., Leamer, E. (Eds.), Handbook of Econometrics, vol. 6B, 5369-5468. Elsevier, Amsterdam.

Imbens, G. 2007. “Instrumental Variables with Treatment Effect Heterogeneity: Local Average Treatment Effects.” NBER Summer Institute 2007. http://www.nber.org/WNE/lect_5_late_fig.pdf. Accessed September 8, 2009.

Imbens, G., Angrist, J. 1994. “Identification and Estimation of Local Average Treatment Effects.” *Econometrica*, 62(2), 467-475.

Imbens, G., Woolridge, J. 2008. “Recent Development in the Econometrics of Program Evaluation.” NBER Working Paper 14251. http://www.nber.org/papers/w14251 Accessed November 2, 2008.

Manski, C. 1990. “Nonparametric Bounds on Treatment Effects.” *The American Economic Review*, 80(2), 319-323.

Manski, C. 1997. “The Mixing Problem in Programme Evaluation.” *The Review of Economic Studies*, 64(4), 537-553.

Mahul, O. 1999. "Optimum Area Yield Crop Insurance." *American Journal of Agricultural Economics*, 81(1), 75-82.

Miquel, R. 2002. "Identification of Dynamic Treatment Effects by Instrumental Variables." University of St. Gallen Discussion Paper 2002-11.

Miranda, M. 1991. "Area-yield Crop Insurance Reconsidered." *American Journal of Agricultural Economics*, 73(2), 233-242.

Moffitt, R. 2008. "Estimating Marginal Treatment Effects in Heterogeneous Populations." Unpublished. http://www.econ.jhu.edu/People/Moffitt/welfls0_v4b.pdf. Accessed August 26, 2009.

Rosenbaum, P. 2002. *Observational Studies*. Springer-Verlag, New York.

Staiger, D., Stock, J. 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica*, 65(3), 557-586.

Train, K. 2009. Discrete Choice Methods with Simulation. Cambridge University Press, New York.

Vercammen, J. 2000. "Constrained Efficient Contracts for Area Yield Crop Insurance." *American Journal of Agricultural Economics*, 82(4), 856-864.