Randomized Encouragement Designs in Agricultural Development: A Methodological Exploration in the Context of Index Insurance

Conner Mullally
UC Davis Development Seminar
May 5, 2010

Abstract: I consider the potential of randomized encouragement designs in agricultural development. The classical problem in program evaluation is the inability to observe individuals both participating and not participating in a program at any point in time. Randomized encouragement designs solve the evaluation problem by randomly varying incentives to participate in a program without affecting outcomes of interest, thus making it possible to measure average treatment effects. They are also of special interest to agricultural development economists for their potential to stimulate program participation; low participation is common in many programs, making it difficult to estimate impacts. These benefits come with caveats that should be taken into account both in research design, and in econometric practice. I explore these issues in the context of evaluating an index insurance product utilizing a randomized encouragement design, mirroring current work in coastal Peru.

Program evaluation is a key concern in agricultural development. Vast resources are expended each year on programs meant to combat poverty and counteract market failures, and it is necessary to estimate the impacts of these programs in order to judge their efficacy. Any program analysis must solve the evaluation problem: How can we estimate the impact of a program if we only observe individuals participating or not participating, and never in the counterfactual state? One option is to compare participants and non-participants, controlling for observable differences and assuming that any unobserved factors are uncorrelated with the outcome or the decision to participate. Alternatively, one may make assumptions about the distribution of unobserved heterogeneity, e.g., that it follows a normal distribution, and control for these differences via functional form. In any case, the evidence strongly suggests that these methods will yield biased estimates of program impacts (Lalonde, 1986; Heckman, 1998). In recent years, applied development economists have followed the lead of their colleagues in labor and public economics, placing emphasis on the use experimental and quasi-experimental methods in program evaluation. Randomized experiments solve the evaluation problem by randomizing treatment, insuring that any unobserved heterogeneity will be uncorrelated with participation, at least in large samples. Quasi-experimental methods are a second-best alternative in lieu of randomized experiments, in which variation in the participation decision that is not correlated with unobserved heterogeneity that may also be driving outcomes is used to explain outcomes of interest. Examples of quasi-experimental methods include instrumental variables (IV), difference-in-differences, and regression discontinuity designs. How closely these methods approximate a randomized experiment will depend on satisfying certain assumptions.

One method that may be implemented via randomization or with observational data is an encouragement design. In this paper, I define an encouragement design as using a variable that changes incentives for program enrollment to predict participation, and using predicted participation to estimate program impacts. It is an application of instrumental variables, and is therefore valid under the same assumptions. I argue that encouragement designs are of particular interest to economists working in agricultural development. While random assignment to treatment is ideal, it may not always be feasible. This may be due to ethical concerns, but low compliance could be an issue as well. In agricultural development, programs may require participants to engage in markets which are unfamiliar, pay some sort of fee to participate, or enter into agreements in which they may not fully understand costs and benefits. These obstacles would discourage participation in any context, but may be especially pronounced in a poor rural economy. Low compliance may result, which will not only have a strong effect on statistical precision, but it will also prevent the kinds of learning processes that are essential to individual participants understanding program benefits. Researchers can address these issues by randomly varying incentives to participate in a way that does not affect outcomes. However, while any properly designed randomized encouragement will yield an unbiased estimate of program effects, the effect will describe impacts on a specific subpopulation. This subpopulation may or may not coincide with the population of primary research interest.

In what follows, these issues are addressed in the context of a randomized encouragement design measuring the impacts of index insurance on decision making by small farmers. This parallels current research in coastal Peru being undertaken by myself and other researchers at UC Davis, along with investigators in Lima. While the context I use is specific, the points raised are relevant for any impact analysis in which takeup of the program and benefits from it enjoyed by individuals are affected by unobserved heterogeneity. Using an economy of risk-averse farmers characterized by mean-standard deviation preferences as an example, I focus on encouraging participation in insurance by varying its price. I illustrate how this research design shapes the effects that may be estimated econometrically, and discuss how to take this into account in research design.

Insuring farmers against systemic risk: index insurance.

Agriculture all over the world is subject to risk. What makes the situation in developing economies different is the absence of well-functioning markets to mitigate the effects of risk. Informal mechanisms exist, but available empirical evidence suggests they are either inadequate or very costly (Rosenzweig and Binswanger, 1993; Dercon, 2004, Carter et al., 2007). Traditional crop insurance, usually based on compensating farmers for losses relative to a farm-specific historical level, has proven extremely expensive in the countries where it is offered. Covering all sources of risk without the ability to perfectly monitor farmer behavior leads to incentive problems for policyholders, and traditional crop insurance has typically seen high ratios of payouts to premium as a result (Skees et al., 1999). This makes traditional insurance a less than desirable policy option for governments in developing countries. Index insurance is an alternative to traditional crop insurance which ties payouts to a variable that is highly correlated with yields or income, but beyond the control of individual farmers. For example, an index insurance contract based on rainfall might pay policyholders if precipitation were to fall below a certain level, with payouts increasing in the size of the shortfall. Since it does not cover all risk at the individual level, index insurance will not offer protection to the same extent as traditional crop insurance, but it avoids the incentive problems which have made such programs costly and unsustainable.

The challenge of impact evaluation and the role of encouragement designs.

While the potential of index insurance as a risk management tool is high, empirical evidence to date is scarce. Studies of demand for index insurance in developing rural economies are few; examples include Giné et al. (2008) and Cole et al. (2009). A small number of studies have gone beyond the question of demand with respect to insurance and developing country agriculture. Giné and Yang (2008) measure the impact of bundling a drought insurance product with a loan on credit demand in Malawi. They reach the surprising conclusion that the offer of insurance decreased demand for credit, and hypothesize that this is due to the limited liability nature of the credit contract in question. Cai et al. (2009) use a cluster-based randomization to measure the impact of a traditional crop insurance product on livestock accumulation in rural China, and find significant and positive effects. To my knowledge, however, no study of the impacts of having index insurance on outcomes of interest has been published to date. Efforts to fill this gap are currently under way, such as our project in Peru and those being led by other researchers in various countries.

Pointing out a lack of evidence begs the question of how to best go about collecting it. As described in the introduction, serious shortcomings with traditional econometric methods in comparison with randomization have been brought to light over the past few decades. However, it is difficult to see how individuals could be randomized to treatment in the context of agricultural insurance. A randomized encouragement design may be a second-best alternative. Examples of such studies in agricultural development are few. Duflo et al. (2009) examine the profitability of fertilizer use among Kenyan maize farmers. Fertilizer is applied to a section of land on each participating farm; the encouragement aspect of the design comes from the fact that the fertilizer is supplied free of charge. Many more examples are found in labor and public economics. Randomly distributed vouchers for payment of school tuition or rent at higher-income housing have been used in several studies to predict education levels and neighborhood choice, which are then in turn used to explain outcomes. See Angrist et al. (2002) for a tuition voucher example, and Katz et al. (2001) for an encouragement design using a housing voucher. In personal finance, Duflo and Saez (2003) study the effects of information on retirement saving decisions, encouraging randomly selected individuals to attend an information session by offering them a small amount of money. A large number of other non-randomized studies can be thought of as encouragement designs as well, exploiting variation in plausibly exogenous factors such as distance or program benefits across US states. For example, McKenzie and Sakho (2010) use distance to a government office to predict formalization of small businesses in Bolivia, and then measure the impact of predicted formalization on profitability. Proximity may cut travel costs or increase the availability of information, encouraging individuals to increase their willingness to formalize.

Beyond solving the evaluation problem, another benefit of any randomized encouragement design from a development perspective is the potential to increase the number of program participants. In the context of index insurance, our experience in Peru and that of some others in different countries (e.g., Giné et al., 2008) has been one of very low takeup rates. On one hand, we might conclude that this is the result of household optimization, and that despite all of its promise, index insurance is not a valuable risk management tool. There are reasons to doubt this conclusion, however. Initiatives of the sort that we are leading in Pisco will often require partnerships with the private sector. In our case, the AYI contract is marketed by La Positiva, a Peruvian insurer. La Positiva had never before offered agricultural insurance, and as a result incurred fixed costs associated with setting up a new division within the company and hiring new personnel. These costs and those of other partners within the project resulted in a premium that was more than double the actuarially fair rate, despite a 30 percent subsidy paid by the Peruvian Ministry of Agriculture. We might not expect this high price to continue as La Positiva recoups its initial outlays for the project, and if the venture proves profitable, when other insurers enter the market. Therefore there some justification for altering the incentives faced by farmers in a way that raises demand, as it may allow us to measure impacts of insurance in an environment that is more representative of what we expect to emerge in the long run. One could imagine similar scenarios in other projects not related to insurance that involve partnerships between research institutions and the business world.

These benefits come with caveats. First and most obvious, identifying mechanisms that effectively change incentives can be difficult. This is where economic theory, baseline data collection, and focus group discussions can serve an invaluable role, enabling researchers to pare down the set of variables that might be manipulated to affect program takeup. Second, as mentioned above in the introduction, the mechanism through which demand is affected will determine the subpopulation to which estimated impacts apply. Suppose there are two different studies of the impacts of AYI on cropping decisions: one that randomizes the insurance premium and another that uses random door-to-door information sessions about the insurance. It stands to reason the sorts of farmers who might be persuaded to buy AYI by the premium discount are different from farmers motivated to buy insurance by the information sessions. In both cases, an instrumental variables estimate of the impacts of insurance would use the randomized encouragement to predict insurance demand, and then use the results of this first stage to estimate impacts of insurance on crop choice. As pointed out repeatedly by Imbens (e.g., Angirst and Imbens, 1995) and Heckman (e.g., Heckman and Vytlacil, 2007), IV estimates of treatment effects only capture average impacts on those individuals who are affected by the instrument; if the instrument does not affect your decision with respect to buying insurance, impacts of insurance on your cropping decision will not factor into the estimated average effect. This makes intuitive sense: IV uses variation in program participation due to the instrument to explain outcomes of interest. Subgroups for which none of the variation in program participation is driven by the instrument will not factor into estimated program impacts. Unless the average impacts of insurance are identical across the group affected by the different encouragements, two researchers applying these research designs to the exact same population will likely emerge with different estimates of the average effect of insurance on crop choice.

In what follows, I extend the example of a randomized price to a two part model of area-yield insurance. In the first part, I examine demand for the insurance among farmers characterized by mean-standard deviation preferences, and show how heterogeneity among the farmers shapes individual and aggregate benefits from AYI. In the second part, I look at this same economy from the perspective of an econometrician charged with carrying out an impact evaluation based on the encouragement design described above. I close with a discussion of how to address the issues raised in the paper in research design.

A simple model of a rural economy subject to risk

The following model draws from Miranda (1991). The economy consists of *N* farmers, each planting a single crop. Crop yield q_i is a random variable with identical mean μ for each of the *N* farmers, and it is affected by a shock ε_c that is common to all *N* households, and an idiosyncratic shock ε_i. Yield for farmer *i* at time *t* is a linear function of these shocks:

$$q_{it} = \mu + \beta_i \varepsilon_{ct} + \varepsilon_{it} \qquad (1)$$

The variables ε_c and ε_i are assumed to be independent and each has a mean of zero. From (1), the variance of output can be decomposed into a component due to the common shock and another due to the household-specific shock:

$$\sigma_{q_i}^2 = \beta_i^2 \sigma_{\varepsilon_c}^2 + \sigma_{\varepsilon_i}^2 \tag{2}$$

Both components of variance have structural parameters that are equal across all households: $\sigma_{\varepsilon_c}^2$ and $\sigma_{\varepsilon_i}^2$. What is different across households is the parameter β_i, which is equal to the covariance of farmer yields q_i with the covariate shock ε_c, divided by the variance of ε_c:

$$\beta_i = \frac{\sigma_{y_i,\varepsilon_c}}{\sigma_{\varepsilon_c}^2} \tag{3}$$

To see this, consider a simple regression model in which output q_i is regressed on the common shock ε_c. The parameter β_i captures the sensitivity of yields for farmer *i* to the covariate shock, and is the source of all heterogeneity between farmers in the economy. At the population level, this is a random variable centered at unity, whereas for an individual farmer it is a parameter whose value is fixed. In the long run, the value of β_i would likely be a choice variable for each farmer, as pointed out by Chambers and Quiggin (2002). In the present context we will assume it is fixed at the individual level.

Each of the *N* farmers in the economy has preferences that may be represented as a function of the mean and standard deviation of random returns, which for now solely consist of crop yields. Normally this would require making assumptions over the distribution of yields or the nature of the ex-post utility function so as to insure that the two-moment representation of preferences is consistent with expected utility or some generalization thereof, as shown by Meyer (1987), Sinn (1983), and Chamberlain (1983). Here, I am less concerned with following the axioms of expected utility than I am with creating a simple, coherent model. In any case, two-moment preference functions often yield close approximation of expected utility results, even when the criteria for consistency with expected utility are not satisfied (Kroll and Markowitz, 1984, Garcia et al., 1994). The preference function is:

$$V(\mu,\sigma) = \mu^2 - \gamma\left(\beta_i^2\sigma_{\varepsilon_c}^2 + \sigma_{\varepsilon_i}^2\right) \tag{4}$$

This function is taken from Nelson and Escalante (2004). It is characterized by constant relative risk aversion; re-scaling random yields q_i will have no effect on the marginal rate of substitution between the mean and standard deviation of returns. In this sense it differs from the usual linear mean-standard deviation specification, which is characterized by constant absolute risk aversion. The parameter γ is the coefficient of relative risk aversion and is assumed to be equal across farmers.

To sharpen the context of the model, I will define the covariate shock ε_c as the deviation of the average yield per hectare in the region from the historical mean. That is:

$$\varepsilon_{ct} = \overline{q_t} - \mu \tag{5}$$

The variation in farmer yields can now be thought of as being decomposed into a component that is explained by variation in area-yields, i.e., average production per hectare, and all other sources risk. Given this definition of the systemic shock, area-yield insurance (AYI) is an obvious risk management tool for farmers in the economy.

An AYI contract will consist of a premium and an indemnity function. Here I choose a simple indemnity function, one that pays farmers the opposite of the covariate shock whenever this shock is negative, i.e., when area-yields fall short of the historical mean, and zero otherwise:

$$I_t = \max[0, -\varepsilon_{ct}] \tag{6}$$

A lower strike point would decrease the absolute value of the covariance between the indemnity and the covariate shock and make the insurance cheaper, but otherwise leave the model unchanged. The premium is equal to the expected indemnity, plus a loading term:

$$p + l = E(I) + l \tag{7}$$

If a farmer decides to buy insurance, her utility is given by:

$$V(\mu,\sigma) = (\mu - l)^2 - \gamma\left(\beta_i^2\sigma_{\varepsilon_c}^2 + \sigma_{\varepsilon_i}^2 + \sigma_I^2 + 2\beta_i\sigma_{\varepsilon_c,I}\right) \tag{8}$$

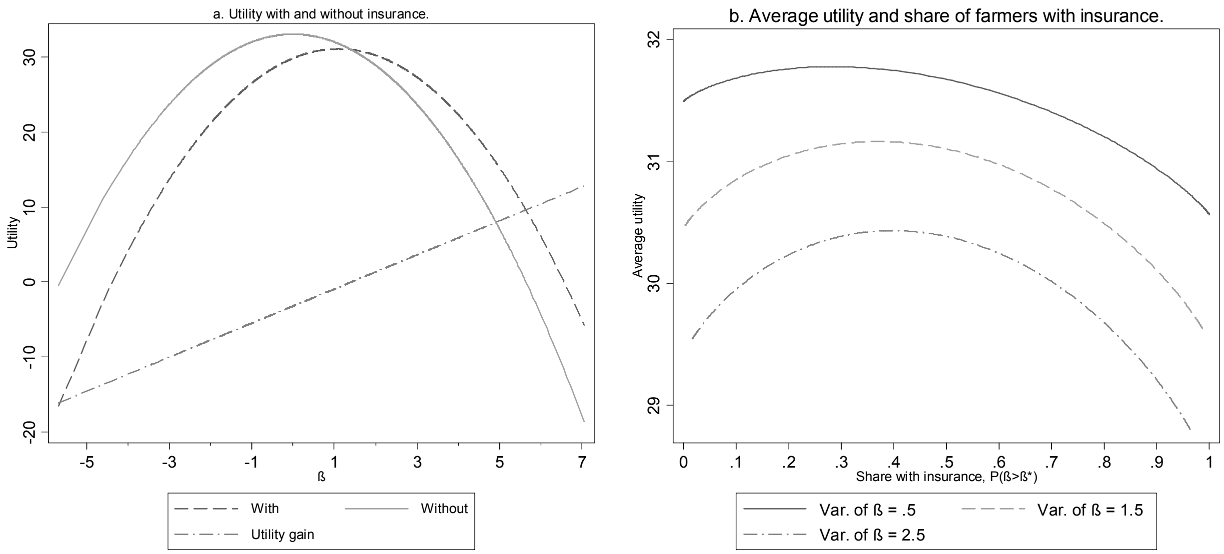
Taking the difference between (8) and (4) yields a decision rule for purchasing insurance. A farmer purchases insurance if her β_i parameter is greater than the critical level β^*, where:

$$\beta^* = \frac{-2\mu l + l^2 - \gamma\sigma_I^2}{2\gamma\sigma_{\varepsilon_c,I}} \tag{9}$$

This is nearly the same decision rule given in Miranda (1991), but differs slightly as his decision rule was based solely on variance reduction rather than utility maximization. Assuming that $\mu > l$ and noting that the denominator is negative, this expression is increasing with respect to the loading term *l*, increasing with respect to the variance of the indemnity σ_I^2, and is decreasing with respect to relative risk aversion γ. It is also increasing with respect to the covariance of the covariate shock and the indemnity $\sigma_{\varepsilon_c,I}$, i.e., as *I* and ε_c become less negatively correlated, β^* increases.

Within the model, the sole source of heterogeneity is the parameter β, and it is the level of this parameter that will determine whether or not a farmer can benefit from AYI. Overall welfare in the economy as a whole will depend on how this parameter is distributed across farmers. This is depicted in the two panels of Figure 1.

Figure 1: Welfare and insurance demand.



Total variance in production is increasing with the absolute value of β_i, and therefore utility without insurance as shown in panel (a) is maximized when this parameter is equal to zero. Utility with insurance has a similar shape, but is everywhere below utility without insurance to the left of β^*, and everywhere above it to the right of β^*. The utility gain from insurance increases linearly with β_i. The slope of the gain line is equal to $-2\gamma\sigma_{\varepsilon_c,I}$, with its intercept at $-2\mu l + l^2 - \gamma\sigma_I^2$.

Panel (b) illustrates how average utility across all farmers changes with the number of farmers purchasing insurance, for a normally distributed β with mean of unity and three different possible values of its variance σ_β^2. The proportion of farmers buying insurance is shifted by varying the loading parameter l in the selection rule, but it is left unchanged in the outcome equation, as if we could discount the insurance at various levels and then measure how well off each farmer would be under the insurance at the regular premium. The analytical expression for this curve is:

$$\overline{V(\mu,\sigma)} = \mu^2 - \gamma\left(\left(\sigma_\beta^2 + 1\right)\sigma_{\varepsilon_c}^2 + \sigma_{\varepsilon_i}^2\right) + P\left(\beta > \beta^*\right)\left(-2\mu l + l^2 - \gamma\left(\sigma_I^2 + 2E\left(\beta \mid \beta > \beta^*\right)\sigma_{\varepsilon_c, I}\right)\right) \quad (10)$$

The first two terms on the right hand side form the expected value of the components that are common to the utility enjoyed by insurance purchasers and non-purchasers alike. The last term in brackets is the expected gain to buying insurance, conditional on $\beta_i > \beta^*$, and weighted by the proportion of the population buying insurance.

I am assuming positive selection, i.e., the first farmers to select into buying insurance are those with the most to gain. As a result, increasing the proportion of farmers with insurance in the economy while holding the distribution of β fixed causes average utility to increase at a decreasing rate and eventually turn downwards, as farmers who are worse relatively off with insurance purchase the index product. In this particular case, β^* is to the right of mean of the normally distributed β (β^*=1.41>1). Widening the spread of β therefore raises the number of farmers for whom $\beta_i > \beta^*$, increasing the proportion of farmers with insurance that maximizes average utility. This is why the probability at which average utility is maximized shifts to the right as I increase the spread of β. The situation would be reversed if, for example, the covariance between the common shock ε_c and the indemnity function were high enough so that β^* was to the left of the mean of β, or if relative risk aversion were sufficiently large. The curvature of the average utility function also increases with the spread of β. To see why, suppose I expand participation in the insurance program by a fixed number of farmers. Given positive selection, farmers brought into the program by the expansion will have lower values of β than farmers already participating. The larger the variance of β, the further away in value will be the β parameters for new and existing participants, and the faster the rate at which the slope of the average utility function depicted in panel (b) will change. In sum, the potential welfare gains due to introducing AYI into this economy hinge on the distribution of β and the parameters of the insurance contract.

Measuring the impacts of index insurance: Average treatment effects

Suppose that the impacts of the AYI contract described above are to be evaluated using a randomized price discount or "coupon," represented by the variable *c*. Given the presence of *c*, the decision rule for purchase of insurance given in (9) becomes:

$$\beta^* = \frac{-2\mu(l - c_i) + (l - c_i)^2 - \gamma\sigma_I^2}{2\gamma\sigma_{\varepsilon_c, I}} \tag{11}$$

The subscript *i* on the variable *c* indicates that this value varies across farmers. Receiving a coupon shifts β^* downward, encouraging farmers to buy insurance.

Assuming that *c* obeys a few assumptions that will be detailed below, it can be used as an instrumental variable in evaluating the AYI contract described in the preceding section.
I will use utility as the outcome of interest. Since I am interested in welfare effects of insurance rather than an insurance subsidy, utility is measured while ignoring the impact of the coupon beyond its influence on purchase of insurance. Direct measurement of utility is an obvious a departure from the reality of program evaluation. Individuals (unfortunately) are not equipped with "utilo-meters" that can be used to gauge welfare. But this setup does reflect the general IV approach to estimation of program impacts: the outcome variable is assumed to be affected by the instrument solely through the latter's effect on program participation. Here, I have created the outcome variable in such a way that this holds by definition.

Now I move from the economic model to an econometric model, based on the information that would typically be available to an econometrician conducting an impact evaluation of the AYI contract described above. First, denote by y_i the outcome as observed by the econometrician and d_i as a dummy variable taking a value of 1 if a farmer buys insurance and 0 otherwise. As a first step, y_i is modeled using a linear random coefficient model:

$$y_i = \lambda_i + \alpha_i d_i \tag{12}$$

The parameter λ_i is equal to utility without insurance in the economic model, $\mu^2 - \gamma\left(\beta_i^2 \sigma_{\varepsilon_c}^2 + \sigma_{\varepsilon_i}^2\right)$, and α_i is the change in utility due to buying insurance, $-2\mu l + l^2 - \gamma\left(\sigma_I^2 + 2\beta_i \sigma_{\varepsilon_c, I}\right)$. The impact on an individual farmer from buying insurance, or treatment effect, is the parameter α_i. Using this model, I can define various average treatment effects. Initially I focus on two: the Average Treatment Effect (ATE), or the average change in utility due to having insurance among all farmers in the population, and the Average Treatment on the Treated (ATT), the average change in utility due to having insurance among all farmers who actually buy it.

Average Treatment Effect (ATE): Average increase in utility for entire population from buying insurance.

$$E(y \mid P(d_i = 1) = 1) - E(y \mid P(d_i = 1) = 0) = E(\alpha) = -2\mu l + l^2 - \gamma\left(\sigma_I^2 + 2\sigma_{\varepsilon_c, I}\right) \tag{13}$$

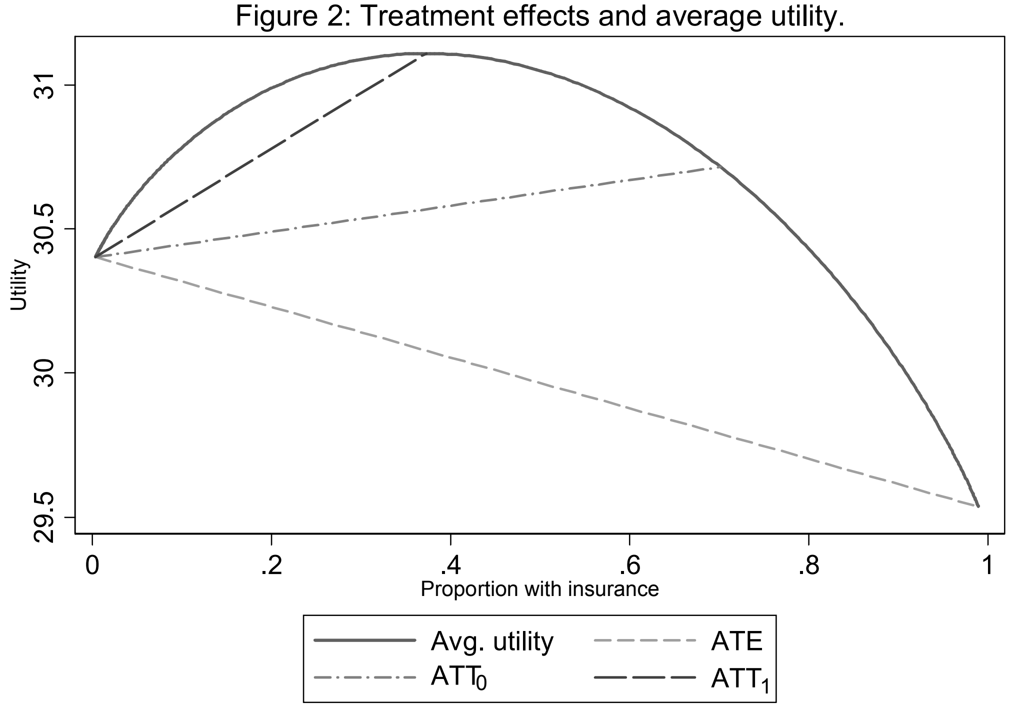
Average Treatment on the Treated (ATT): Average increase in utility due to buying insurance for those who actually buy it.

$$\begin{aligned} &E(y \mid P(d_i = 1) = p) - E(y \mid P(d_i = 1) = 0) = E(\alpha \mid P(d_i = 1) = p) = \\ &-2\mu l + l^2 - \gamma\sigma_I^2 - 2\gamma E\left(\beta_i \mid \beta_i > \beta^*, P(\beta_i > \beta^*) = p\right)\sigma_{\varepsilon_c, I} \end{aligned} \tag{14}$$

for some $0<p<1$. The expression $E\left(\beta_i \mid \beta_i > \beta^*, P(\beta_i > \beta^*) = p\right)$ is the expected value of β taken over the values of β for all farmers who participate (i.e., all those for whom $\beta_i > \beta^*$) when the proportion of farmers participating is equal to p. $P(d_i = 1)$ is the proportion of farmers buying insurance, and is equal to $P(\beta_i > \beta^*)$ in the economic model. These effects capture average impacts assuming that outcomes of a given individual are not affected by whether other individuals purchase insurance, i.e., no general equilibrium effects. The ATT can be evaluated at various values of p, not all of which are equally relevant from the perspective of policy evaluation. Ideally, one could estimate the ATT at various values of p, to get an idea of how the gains to insurance change with the proportion of farmers electing to buy AYI. In any case, the average treatment effects that are estimated will depend on what I can econometrically identify, and may or may not correspond to (13) or (14).[1]

The ATE and ATT can be represented graphically using the average utility curve shown in panel (b) of Figure 1:

Figure 2: Treatment effects and average utility.



The slopes of ATE, ATT_0, and ATT_1 as depicted in the figure are the respective treatment effects. The probability of takeup in the figure is manipulated along the x axis by changing the

[1] Note that I will sometimes talk of an "average treatment effect" when referring to the average impact of insurance on utility for some sub-population. This is not be confused with the ATE, which will always refer to the average change in utility that occurs when everyone in the population has insurance.

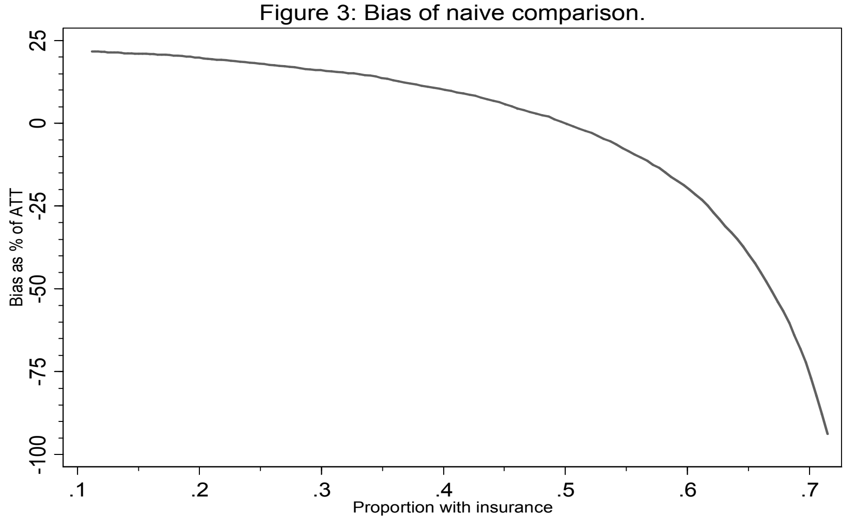
value of the coupon c, thereby shifting β^*. Since β^* in the absence of the subsidy is greater than the median of β, a majority of farmers are made worse off by purchasing insurance, and as a result the ATE slopes downwards. This situation could be reversed by a smaller loading factor l, for example. ATT_0 is evaluated at the actuarially fair insurance contract, i.e., $c=l$. ATT_1 is evaluated at the proportion of farmers electing to buy insurance in the absence of the subsidy, $c=0$. Due to positive selection, the ATT falls as the proportion of farmers buying insurance expands. Those with the most to gain from buying insurance elect to purchase it when the price is high, driving up average gains to buying insurance at low values of c.

Suppose the impact of AYI were estimated using a comparison between purchasers and non-purchasers. This would yield an estimate of the following:

$$\begin{aligned} &E(\lambda+\alpha \mid d_i=1)-E(\lambda+\alpha \mid d_i=0)= \\ &\gamma\sigma_{\varepsilon_c}^2\left(\left(\sigma_\beta^2 \mid d_i=0\right)-\left(\sigma_\beta^2 \mid d_i=1\right)\right)+l^2-2\mu l-\gamma\left(\sigma_I^2+2E(\beta \mid d_i=1)\sigma_{\varepsilon_c,I}\right) \end{aligned} \tag{15}$$

where $\left(\sigma_\beta^2 \mid d_i=0\right)$ is the variance of β given that $\beta<\beta^*$. The sign and magnitude of the bias therefore depends on the value of β^*. Figure 3 graphs the bias of this estimate as a proportion of the ATT:

Figure 3: Bias of naive comparison.



The bias approaches zero as β^* goes to unity; since the distribution of β is symmetric, splitting the distribution at $\beta = 1$ yields two half-normal distributions with equal variance, and the first term in the second line of (15) drops out. Where the graph dips below zero the sign and magnitude of the effect estimated in (15) are both incorrect.

Now suppose c is used as an instrumental variable. For a given farmer, c_i can take on one of J distinct values, i.e., $c_i \in [c^1, c^2, \ldots c^J]$. Call this set of possible values C. Rewrite the model given in equation (12) as:

$$y_i = \lambda_i + \alpha_i d_i \tag{16}$$

$$d_i = 1\{\omega_i c_i + e_i \geq 0\} \tag{17}$$

The model now includes a selection equation, (17), which models the decision to buy insurance as being additively separable in the observable determinant of the decision to purchase insurance, c_i, and the unobservable components of the decision, represented by e_i. The impact of the coupon on the purchase decision is captured by the parameter ω_i. Both e_i and ω_i are fixed for a given farmer i and follow some unknown joint distribution within the population. Taking expected values of the model conditional on the value of c yields:

$$E(y \mid c_i = c) = E(\gamma \mid c_i = c) + E(\alpha \mid d_i = 1, c_i = c)E(d \mid c_i = c) \tag{18}$$

$$E(d \mid c_i = c) = P(\omega_i c_i + e_i \geq 0 \mid c_i = c) = P(d = 1 \mid c_i = c) \tag{19}$$

The expectation of (19) is the "propensity score," i.e., the probability that a farmer with a given value of the observable characteristic c decides to buy insurance. Manipulating the propensity score by varying c allows us to measure how the expected outcome would change if one were to vary the proportion of farmers buying insurance within the population.

Equations (18) and (19) suggest *c* could be used to estimate average treatment effects. In Figure 2, the different treatment effects are given by the slopes of the different line segments, i.e., the change in average utility over the change in the probability of participation. If the coupons are shifting average utility and the probability of purchasing insurance, the conditional expectations given in (18) and (19) can be used to construct parameters of the form:

$$\frac{E(y \mid c_i = c'') - E(y \mid c_i = c')}{P(d_i = 1 \mid c_i = c'') - P(d_i = 1 \mid c_i = c')} \tag{20}$$

This is the slope coefficient for a line segment connecting two points on the average utility curve in Figure 2. Any pair of coupon values $(c'', c') \in C$ such that $c'' \geq c'$ can be used to construct a parameter of this form. Whether the resulting slope coefficients represent average treatment effects of insurance on utility depends on whether *c* satisfies several assumptions. Those assumptions are:

$$E(\lambda \mid c_i = c) = \overline{\lambda} \textbf{ for all } c \in C \tag{21}$$

$$E(\alpha \mid c_i = c) = g(P(d_i = 1 \mid c_i = c)) \textbf{ for all } c \in C \tag{22}$$

$$\sigma_{c,d} \neq 0 \tag{23}$$

Denote by $d_i^{c''}$ the value of d_i when $c_i = c''$. Then $d_i^{c''} - d_i^{c'} \geq 0$ for all *i* and all pairs $(c'', c') \in C$ such that $c'' \geq c'$. (24)

Assumption (21) states that the instrument is mean independent of the outcome in the absence of insurance. This rules out the possibility, for example, that a group of farmers receiving a higher coupon value than their counterparts would have higher or lower average outcomes without insurance than farmers receiving smaller coupons. Assumption (22) states that the average outcome with insurance among farmers is a function of the proportion of farmers that decide to buy insurance. Average utility increases or decreases based on the proportion of farmers electing to purchase insurance, and not because of the size of the coupon. Assumption (23) states that the covariance between the decision to buy insurance and the coupon variable cannot be equal to zero; the instrument must have some predictive power. Finally, assumption (24) is the "monotonicity" assumption first proposed by Imbens and Angrist (1994). All farmers that would buy insurance if they were to receive a coupon with value c' would also elect to buy the insurance if they received a larger coupon c''. Note that this relationship could work in the opposite direction, i.e., that everyone that would select into the program if chosen to receive a larger coupon would also select in if chosen to receive a smaller coupon. What matters is that it must hold for everyone in order for econometric estimates based on instrumental variables to represent an average treatment effect of insurance for at least some members of the population. Note that this assumption is equivalent to assuming that $\omega_i \geq 0$ for all i in (17). Given a scalar instrument, monotonicity implies and is implied by the additively separable structure given in (17), as shown by Vytlacil (2002).

Given these assumptions, for every pair of coupon values c' and c'', one can estimate an average treatment effect of the following form given in (20). In terms of the economic model, this effect is:

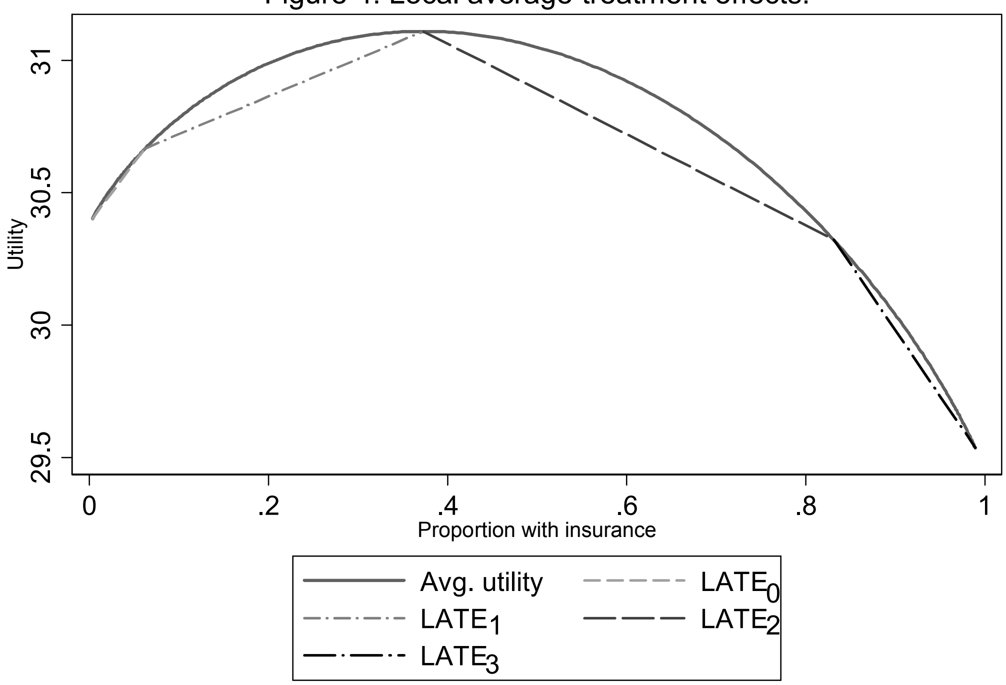
$$-2\mu l + l^2 - \gamma\sigma_I^2 - \left(\frac{P(\beta_i > \beta^*(c''))(2\gamma E(\beta_i \mid \beta_i > \beta^*(c''))\sigma_{\varepsilon_c, I}) - P(\beta_i > \beta^*(c'))(2\gamma E(\beta_i \mid \beta_i > \beta^*(c'))\sigma_{\varepsilon_c, I})}{P(\beta_i > \beta^*(c'')) - P(\beta_i > \beta^*(c'))} \right) \quad (25)$$

where $\beta^*(c'')$ represents equation (11) evaluated at the coupon value c''. This is a Local Average Treatment Effect (LATE). Given positive selection into insurance and assumptions (21) through (24), the LATE can be interpreted as what the change in average utility among the population would be if we were to shift the proportion of farmers with insurance from

$P(\beta_i > \beta^*(c'))$ to $P(\beta_i > \beta^*(c''))$. This gain in average utility comes from farmers induced to buy insurance by receiving a coupon of size c'' rather c'. Thus the estimated LATE is informative about gains from insurance enjoyed by farmers falling within a specific portion of the β distribution, and the size of this portion depends on the shift in participation probability generated by the moving the instrument from c' to c''. Note that if no farmers buy insurance when receiving a coupon of c', the LATE becomes the ATT evaluated at $P(d_i = 1 \mid c_i = c'')$.

Figure 4 shows LATEs evaluated at several different pairs of values for $P(d_i = 1)$, where the variation in $P(d_i = 1)$ was generated using equally spaced values of *c*:

Figure 4: Local average treatment effects.

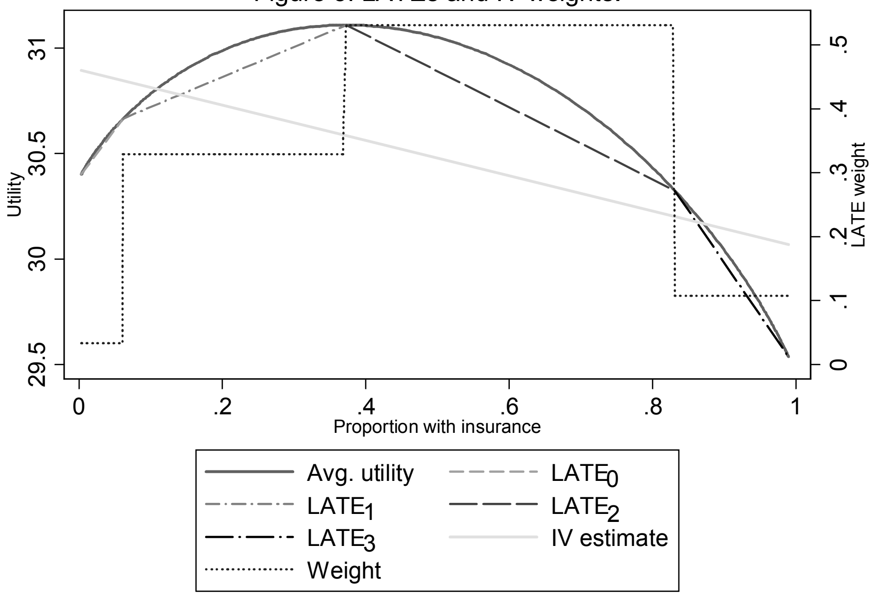


Here each LATE corresponds to two neighboring coupons, but a LATE can be estimated for any two coupons in the set *C*. LATEs can be positive or negative; the instrument *c* generates variation in the purchase decision, and the sign of the LATE will depend on where this variation is located in the support of $P(d_i = 1)$. One could also use all of the values of c to estimate a single average effect via two-stage least squares (for example) rather than a series of LATE estimates, using $P(d_i = 1 \mid c_i = c)$ as an instrument for d_i. The result would be a weighted average of the pairwise LATEs. Denote $P(d_i = 1 \mid c_i = c^l)$ by p_l. The weight assigned by two-stage least squares to the LATE that connects the values of average utility at p_l and p_{l+1} is:

$$(p_{l+1} - p_l)\frac{\sum_{t=l+1}^{K}(p_t - E(P(d_i = 1)))\pi_t}{\sigma^2_{P(d_i=1)}} \tag{26}$$

where π_t is the proportion of the sample for which $P(d_i = 1) = p_t$, or equivalently, the proportion for which $c_i = c^t$. The parameter $\sigma^2_{P(d_i=1)}$ is the variance of $P(d_i = 1)$, and $E(P(d_i = 1))$ is the expected value of $P(d_i = 1)$ taken over all values of c. These weights were first derived by Yitzaki (1996) and applied to the treatment effects model by Heckman and Vytlacil (2007). Note that they are a function the shift in the probability of participation generated by the instrument for each LATE. Here I can calculate the weights exactly, and in real-world applications they can be estimated from data. Figure 5 adds the two-stage least squares estimate of the average impact of having insurance and the weights given in (26) for each LATE to Figure 4:

Figure 5: LATEs and IV weights.



The greatest shift in the probability of participation is observed for $LATE_2$, and it receives the most weight in the two-stage least square estimate; the latter is depicted by the downward sloping light grey line. The LATEs remain positive up to 40 percent participation in the insurance program, but the weighted average of these effects is negative. This is not to say that the ATE and ATT cannot be estimated; in the example given in Figure 5, the ATE could be estimated as a LATE using the smallest and largest values of the coupon, and various ATT parameters could be estimated using the smallest value of the coupon. But it is important to note that while combining all values of the coupon and estimate an average effect via two-stage least squares can increase precision, this effect will not generally correspond to any of the usual treatment parameters, and is not easy to interpret.

Average treatment effects and research design

In the presence of treatment effect heterogeneity, the interpretation of estimated average treatment effects will depend on where the variation generated by the instrument falls on the support of $P(d_i = 1)$. Under positive selection into the program, we can say something about the gains to insurance among those affected by the instrument relative to the population in general. If variation generated by the instrument is located near the upper boundary of $P(d_i = 1)$, estimated LATEs will likely understate the ATE, and certainly any ATT.

This has clear implications for interpretation of estimated effects, as well as for research design. Consider the case of power analysis. Research resources are scarce, and because of this, the potential of a research design to demonstrate impacts of the program being investigated must be taken into consideration. This involves conducting a power analysis in order to estimate "minimum detectable effects." A minimum detectable effect (MDE) is the smallest possible effect of a program that, if true, has an X% chance of producing an impact estimate which rejects the null hypothesis of no effect at a given level of significance, where X is usually set equal to 80 (Bloom, 1995). Suppose the research design includes an encouraged group, which is assigned a value of the instrument equal to c'', and a control group, which receives c'. Then the MDE formula is:

$$MDE = 2.49 * \sqrt{\left(\frac{\sigma^2(1-R^2)}{N\pi_{c''}\pi_{c'}}\right)} \frac{1}{P(d_i = 1 \mid c = c'') - P(d_i = 1 \mid c = c')} \qquad (27)$$

where σ^2 is the variance of the outcome, R^2 is the goodness-of-fit from a regression of the outcome on the treatment indicator and whatever other covariates are available, N is the sample size, $\pi_{c''}$ is the proportion of the sample assigned $c = c''$, and the propensity scores are defined as before. This is the formula for the MDE that has an 80% of rejecting the null of no effect under a one-tailed test with a 5% significance level, and it lies 2.49 standard deviations to the right of zero in the standard normal distribution (see Bloom for an explanation of the 2.49 figure).

If there is no treatment effect heterogeneity, then the only likely implication of having a value for the difference in propensity scores that is less than unity will be a loss of precision.[2] But if there is treatment effect heterogeneity, then not only will the size of the MDE grow, but the underlying true effect on compliers will change as well. The difference in the propensity scores at the two values of the instrument tells us where the estimated effect will be located on the horizontal axis of Figure 5. In the case of positive selection, increasing the participation rate among the encouraged group decreases the size of the true average effect that will be estimated from data, holding the takeup rate among the control group fixed. Given the concave shape of the average response curve generated by positive selection, the true average effect may also become negative as participation rates change in this way.

Holding the participation rate of the encouraged group fixed while increasing that of the control group has the same impact on the true effect. Any MDE may eventually become larger than the true average effect if participation rates of the encouraged and discouraged groups are shifted as described above. The true average effect may also change sign, if the encouraged group is actually worse off when treated, based on average changes in the outcome of interest. Similarly, for a fixed difference in participation rates between the control and encouraged groups, moving the compliance rate among encouraged subjects closer to 100 percent will decrease the true average impact of the program on compliers, making it more likely that a given MDE is greater than the true average effect of the program on compliers, and may eventually change the sign of the true average effect.

These considerations can influence the choice between research designs that might otherwise appear equal. Suppose a researcher is considering two different designs: a

[2] If the difference is very small, then the use of IV may no longer be justified, regardless of exogeneity of the instrument (Bound et al., 1995; Staiger and Stock, 1997).

randomization of eligibility, and a coupon scheme that allows anyone to buy the insurance, but randomly varies the premium across farmers. Suppose further that our researcher expects differences in takeup rates between encouraged and control group farmers to be equal in the two schemes. A standard MDE analysis would suggest that the design with the lowest cost is the best choice. If impacts are likely to be heterogeneous and that the program is characterized by positive selection, the first design has several advantages. First and most obviously, this is because it would yield an estimate of the ATT rather than a LATE with a more opaque interpretation. But secondly, under positive selection the true effect on compliers will be larger under the first design, making rejection of the null hypothesis of no effect more likely. The second design may also place the variation in the takeup rate in a section of the average utility curve where the outcome of interest is decreasing on average. This would virtually eliminate the possibility of estimating a positive average treatment effect altogether.

We could make educated guesses about how large the difference in the true effects under the two competing designs might be, if we want to assume a model for the unobserved heterogeneity that will influence program participation, such as the model combining positive selection and the normal distribution used above. Baseline data can be used to characterize the degree of heterogeneity across observations. In the above example of index insurance, the ideal would be to compare movements in average yields with more disaggregated levels of production, to get an idea of the spread of the β distribution. Detailed time series data of this sort are usually lacking in developing countries. In lieu of such data, one could speculate as to what factors might contribute to the sensitivity of output to covariate risk, and examine whatever data are available that might capture the similarity of these variables across the population of interest. For example, suppose agriculture in the study region is highly dependent upon irrigation. Heterogeneity in sensitivity to covariate risk could be driven by distance of land parcels from canals. These sorts of data might be maintained by local agricultural extension offices, and if not, are easily collected using GPS devices.

More important than considerations of power analysis are the links between choosing the effects of insurance we are most interested in learning about, and designing the instrument that will be used to predict demand. Consider the choice between two coupon-based research designs: one that will split subjects into two or three groups and randomly allocate coupons of different sizes, and another that will turn the coupon into a continuous instrument. An obvious benefit of the first design is that it is simpler, and by choosing which few coupon values to allocate, can also be made less costly than the second design. The first design will allow us to estimate takeup rates at the different coupon values using very few assumptions, and construct estimated LATEs accordingly. Given knowledge of how farmers will respond to the different incentives, coupon values can be chosen to yield estimates of the ATT at different takeup rates. These ATT estimates can be used in combination with data on program costs to conduct cost-benefit analysis, comparing total costs of the program to the total benefits enjoyed by those participating in the program. If little is known about the shifts in the probability of buying insurance that will be generated by the encouragement, then this research design may be enough to show a positive effect on some group. In this case, the points made above with respect to MDEs become relevant, and the research design should aim to generate variation in takeup rates via the instrument where they are most likely to yield significant estimates of average treatment effects.

Perhaps the main disadvantage of this second design is that it does not allow for learning much about the shape of the response curve, or in the index insurance example, average utility curve, as depicted in the various figures above. At best it will yield a rough idea of whether it is convex or concave. It is also difficult to construct more sophisticated estimates of the gains from the program, beyond the simple cost-benefit measure described above. For example, suppose we are interested in the efficiency of subsidizing the purchase of AYI, such as we have done with our coupon design. In an efficient policy, the change in expected costs caused by bringing one more person into the program will be equal to the expected benefits of doing so. A complete measure of efficiency would require knowledge of the marginal benefit and costs of all parties involved. However, the expected gains of marginal program participants and the marginal cost of expanding participation can be compared using the instrument. Using the research design that includes a continuous instrument, the average response curve could be estimated using a flexible functional form, such as a polynomial function of the propensity score. The change in expected benefits is equal to the derivative of my estimated response function, which can be evaluated at any point on the support of the propensity score. This is a continuous version of the LATE known as the Marginal Treatment Effect in the program evaluation literature (Bjorklund and Moffitt, 1987; Heckman and Vytlacil, 2007). This would be compared with change in average cost induced by expanding participation at this same point on the support of the propensity score, which is equal to the value of the coupon at that point.

Under a design with a continuous instrument, all of the average treatment effect parameters available to us under the simpler design can be estimated, although with less precision. Under the simple design, the components of an estimated LATE consist of four different sample averages. As the number of observations at each coupon value shrinks, this method becomes less precise. Heckman and Vytlacil (2007) prove that the LATE and all other average treatment effects can be expressed as weighted averages of MTEs. Estimating average treatment effects as weighted averages of MTEs is a highly data intensive process, as the weights and MTEs must all be estimated, and a loss of precision relative to estimation under the simpler design can result. If the instrument is discrete and multi-valued, then weighted averages of LATEs can be estimated via two-stage least squares, as shown in Figure 5. Estimates such as these can be used to show that there is a positive effect on some subpopulation, but are difficult to interpret. Alternatively, one could estimate the components of (20) non-parametrically, e.g., using local regression. If the support of the propensity score includes 0 and 1, the ATE and various ATTs can be estimated in this way. This strategy is subject to the usual limitations of nonparametric methods, e.g., tradeoffs between bias and variance. The more complex research design offers greater flexibility, and is ideal for learning about heterogeneity in treatment effects. But there are risks of losing precision, ease of interpretation, and the obvious difficulties associated with complicating any research program. Researchers therefore need to think hard about which sorts of effects are of greatest interest, and design the randomized encouragement accordingly.

Answering detailed questions about the pros and cons of different research designs will require some knowledge of how takeup rates are likely to respond to different values of the instrument, and this will likely require baseline data collection. In the case of index insurance, this could be done by offering farmers a menu of increasingly larger hypothetical premiums, and finding the point at which each is unwilling to pay for insurance. The literature from experimental economics has shown repeatedly that hypothetical questions tend to elicit biased responses, as subjects tend to overstate willingness to pay. Blumenschein et al. (2008) offer some suggestions so as to minimize this bias. For example, rather than giving subjects a Yes/No decision at each price, each could be asked if she were “very willing to pay,” “somewhat willing to pay,” or “not at all willing to pay” at each price. Only the “very willing to pay” responses could be counted as a sign of willingness to pay at the indicated price. These methods are certainly not perfect, but they will give us some information where previously there was none. Experimental methods that allow subjects to make choices with real economic incentives may also be of use. Galarza (2009) conducts an experiment in coastal Peru in which participants choose between three contracts, one of which simulates an AYI product, in order to measure the impact of introducing index insurance on demand for credit and willingness to make risky investments. Just as is the case with hypothetical valuation methods, experiments will not give us an ideal picture of willingness to pay, as making payoffs equal to their real world counterparts would be prohibitively expensive. But this is far superior to the alternative of total ignorance.

An issue I have not touched upon as yet is the ability of farmers to understand a program as sophisticated as index insurance. This will obviously have important ramifications for the ability of any instrument to affect participation, and for the potential of insurance or any other program to generate impacts on outcomes of interest. If farmers are very unsure about the benefits gleaned from buying AYI, then they may be willing to buy it conditional on receiving a price discount of sufficient size, but they cannot be expected them to change behavior in terms of measurable outcomes of interest. This effectively adds another dimension to unobserved heterogeneity, as the participation decision will be affected by potential gains to insurance, the ability to evaluate the contract, and possibly the interaction of the two. This is another area where baseline data collection can be extremely valuable. At a most basic level, data on average education levels and experience with the formal financial sector can be used to guess whether a typical farmer would be able to understand an insurance contract. More detailed questions about financial literacy or the ability to understand insurance contracts can also be included. The effects of different kinds of information on the ability to understand sophisticated contracts such as index insurance can be measured by asking financial literacy questions before and after exposing surveyed farmers to information about the index insurance product. This is clearly a large area of research in itself, and I leave a detailed treatment of these issues to future work.

Economists working in agricultural development are not only increasing their focus on program evaluation, but are also becoming involved in the earliest stages of research design. In this paper I have examined a randomized encouragement design in the context of index insurance, an issue at the forefront of agricultural development policy. Randomized encouragement designs are useful not only for their ability to solve the evaluation problem, but for the effect they have on stimulating program participation. These benefits come with caveats, however, and researchers have hard choices to make with respect to research design. Specifically, researchers must consider the following:

1. What are the average treatment effects that are of greatest interest?
2. How important is it to capture heterogeneity in treatment effects?
3. Given the size of the sample and what can be learned about likely treatment effect heterogeneity from baseline data, what are the likely tradeoffs between (1) and (2)?

Once these questions have been answered, researchers should pick the simplest research design that makes it possible to achieve research priorities.

References

Angrist, J., Bettinger, E., Bloom, E., Kling, E., Kremer, M. 2002. "Vouchers for private schooling: Evidence from a randomized natural experiment." *American Economic Review*, 92(5), 1535-1558.

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.

Bloom, H. 1995. "Minimum detectable effects: A simple way to report the statistical power of experimental designs." *Evaluation Review*, 19(5), 547-556.

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.

Carter, M., Little, P., Mogues, T., Negatu, W. 2007. "Poverty traps and natural disasters in Ethiopia and Honduras." *World Development*, 35(5), 835-856.

Chamberlain, G. 1983. "A characterization of the distributions that imply mean-variance utility functions." *Journal of Economic Theory*, 29(1), 185-201.

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.

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

Cole, S., Giné, X., Tobacman, J., Topalova, P., Townsend, R., Vickery, J. 2009. "Barriers to household risk management: evidence from India." Harvard Business School Finance Working Paper 09-116. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1374076.

Dercon, S. 2004. "Growth and shocks: evidence from rural Ethiopia." *Journal of Development Economics*, 74(2), 309-329.

Duflo, E., Kremer, M., Robinson, J. 2008. "How high are rates of return to fertilizer? Evidence from field experiments in Kenya." *American Economic Review*, Papers and Proceedings, 98(2), 482-488.

Duflo, E., Saez, E. 2003. "The role of information and social interactions in retirements plan decision: evidence from a randomized experiment." *Quarterly Journal of Economics*, 118(3), 815-842.

Garcia, P., Adam, B.D., Hauser, R.J. 1994. "The use of mean-variance for commodity futures and options hedging decisions." *Journal of Agricultural and Resource Economics*, 19(1), 32-45.

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

Giné, X., Yang, D. 2008. "Insurance, credit, and technology adoption: Field experimental evidence from Malawi." *Journal of Development Economics*, 89(1), 1-11.

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.

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.

Hanson, S.D., Ladd, G.W. 1991. "Robustness of the mean-variance model with truncated probability distributions." *American Journal of Agricultural Economics*, 73(2), 436-445.

Heckman, J., Ichimura, H., Smith, J., Todd, P. 1998. "Characterizing selection bias using experimental data." *Econometrica*, 66(5), 1017-1098.

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.

Imbens, G., Angrist, J. 1994. "Identification and estimation of local average treatment effects." *Econometrica*, 62(2), 467-475.

Katz, L., Kling, J., Liebman, J. 2001. "Moving to opportunity in Boston: Early results of a randomized mobility experiment." *Quarterly Journal of Economics*, 116(2), 607-654.

Kroll, Y., Levy, H., Markowitz, H.M. 1984. "Mean-variance versus direct utility maximization." *The Journal of Finance*, 39(1), 47-61.

Lalonde, R. 1986. "Evaluating the econometric evaluations of training programs with experimental data." *American Economic Review*, 76(4), 604-620.

McKenzie, D., Sakho, Y. 2010. "Does it pay firms to register for taxes? The impact of formality on firm profitability." *Journal of Development Economics*, 91(1), 15-24.

Meyer, J. 1987. "Two-moment decision models and expected utility maximization." *American Economic Review*, 77(3), 421-430.

Miranda, M. 1991. "Area-yield crop insurance reconsidered." *American Journal of Agricultural Economics*, 73(2), 233-242.

Nelson, C.H., Escalante, C. 2004. "Toward exploring the location-scale condition: a constant relative risk aversion location-scale objective function." *European Review of Agricultural Economics*, 31(3), 273-287.

Rosenzweig, M., Binswanger, H. 1993. "Wealth, weather risk and the composition and profitability of agricultural investments." *Economic Journal*, 103(416), 56-78.

Sinn, H.W. 1983. *Economic Decisions under Uncertainty*. New York: North Holland Publishing.

Skees, J., Hazell, P., Miranda, M. 1999. "New approaches to crop yield insurance in developing countries." International Food Policy Research Institute EPTD Discussion Paper 55. http://www.ifpri.org/sites/default/files/publications/eptdp55.pdf.

Vytlacil, E. 2002. "Independence, monotonicity, and the latent index models: An equivalence result." *Econometrica*, 70(1), 331-341.

Yitzhaki, S. 1996. "On using linear regression in welfare economics." *Journal of Business & Economics Statistics*, 14(4), 478-486.