Taking a New Look at Empirical Models of Adoption: Average Treatment Effect Estimation of Adoption Rates and their Determinants

Aliou Diagne
Impact Assessment Economist
The Africa Rice Center (WARDA)
B.P. 2031 Cotonou, Benin
Tel: (229) 21 35 01 88 Extension: 333
Fax: (229) 350 556
E-mail: A.Diagne@cgiar.org

Contributed paper prepared for presentation at the International Association of Agricultural Economists Conference, Gold Coast, Australia, August 12-18, 2006

Taking a New Look at Empirical Models of Adoption: Average Treatment Effect Estimation of Adoption Rates and their Determinants

1. Introduction

The paper approaches the problem of estimation of adoption rates and their determinants from the perspective of modern evaluation theory as exposed in the treatment effect estimation literatures (for a review of the literatures, see Wooldridge, 2002; Blundell and Costa Dias, 2000; Heckman, et. al., 1999; Angrist et. al., 1996; Moffit, 1991). As demonstrated below, this approach is necessary because commonly used adoption rates estimators suffer either from what we call "non-exposure" bias or from selection bias. As a consequence, they generally yield biased and inconsistent estimates of *population* adoption rates even when based on a randomly selected sample. The "non-exposure" bias results from the fact that farmers who have not been exposed to a new technology cannot adopt it *even if they might have done so if they had known about it*. This results in the *population* adoption rate being underestimated.

Because of the "non-exposure" bias resulting from the incomplete diffusion of the technology in the population, the *observed* proportion of sample farmers who have adopted does not consistently estimate the true population adoption rate even with a random sample. The sample adoption rate within the sub-sample of farmers exposed to the technology is *not* a consistent estimate of the true population adoption rate either (even if the sample is random). Likewise, the effects of determinants of adoption cannot be consistently estimated using a simple probit, logit or Tobit adoption model that does not control for exposure. This difficulty in interpreting the coefficients of the simple probit, logit or Tobit adoption model when the diffusion of the technology in the population is not complete has been pointed out by Besley and Case (1993) Saha et. al., (1994) and Dimara and Skuras (2003).

It turns out that the population adoption rate corresponds to what is defined in the modern treatment effect literature as the *average treatment effect,* commonly denoted by *ATE*. ATE measures the effect or impact of a "treatment" on a person randomly selected in the population. In the adoption context, a "treatment" corresponds to exposure to a technology and the average treatment effect on the adoption outcomes of population members is the (potential) *population* adoption rate. That is, the adoption rate when *all* members of the population have been exposed to the technology. Another quantity that is also the subject of attention in the treatment effect literature is the *average treatment effect on the treated,* which is commonly denoted by *ATE1*. The ATE1 is a measure of the effect of treatment in the treated subpopulation and corresponds in the adoption context to the adoption rate among the exposed. [1]

The paper is organized as follows. Section 2 uses the counterfactual outcomes and *average treatment effect* estimation framework to derive consistent non-parametric and parametric estimators of population adoption rates and their determinants. Section 3 applies the results of section 2 to consistently estimate the population adoption rates and determinants of the NERICA (New Rice for Africa) rice varieties in Cote d'Ivoire along with estimates of the population "adoption gap" and selection biases created by the presently limited diffusion of the NERICAs. Section 4 concludes the paper with a summary of the major methodological and empirical results of the paper and their policy implications.

2. ATE estimation of population adoption rates and their determinants.

Following the modern treatment effect estimation literature (see, for example, Wooldridge, 2002; Heckman, 1996; Angrist et. al., 1996; Rosenbaum and Rubin, 1983), we use

[1] The two concepts of adoption and diffusion are often used interchangeably in the voluminous adoption literature (see Feder et al. 1985 and Sunding and Zilberman, 2001, for reviews of the literature). The adoption of a technology is defined in this paper to mean its *use* at the individual farmer level or at the aggregate population level. The term diffusion is used strictly in this paper to mean the extent of *awarness* of (or *exposure* to) the technology in the population (which does not necessarily implies its use)

a *counterfactual* outcome framework by which every farmer in the population has two *potential* outcomes: with and without exposure to a technology. For concreteness and without loss of generality, we will focus on the adoption of new varieties. Let y_1 be the potential adoption outcome of a farmer when exposed to the new varieties and y_0 be the potential adoption outcome when not exposed to them. The potential adoption outcome can be either adoption status (a dichotomous 0-1 variable) or a measure of intensity of adoption such as the total land area allocated to the new varieties.[2] Then, the "treatment effect" for farmer *i* is measured by the difference $y_{i1} - y_{i0}$. Hence, the expected population adoption impact of exposure to the new varieties is given by the expected value $E(y_1 - y_0)$, which is, by definition, *the average treatment effect*, ATE.[3]

But, the inability to observe both an outcome and its counterfactual makes it impossible in general to measure $y_1 - y_0$ for any given farmer. However, since exposure to a new variety is a necessary condition for its adoption, we have $y_0 = 0$ for any farmer whether exposed to the set of new varieties or not. Hence the adoption impact of a farmer *i* is given by y_{i1} and the average adoption impact is given by $ATE = Ey_1$. Unfortunately, we observe y_1 only for farmers exposed to the new varieties. Hence, we cannot estimate the expected value of y_1 by the sample average of a randomly drawn sample since some of the y_1 in the sample would be missing.

If we let the binary variable *w* be an indicator for exposure to the varieties, where *w* =1 denotes exposure and *w* =0 otherwise. The average adoption impact on the exposed subpopulation is given by the conditional expected value $E(y_1 \mid w = 1)$, which is by definition the

[2] We will focus on adoption status in the empirical analysis, however.

[3] Here y_1 and y_0 are considered random variables, representing the potential outcomes of any farmer randomly selected from the underlying population of farmers. Similarly, *y* will denote the *observed* adoption outcome of a randomly selected farmer.

average treatment effect on the treated, commonly denoted by ATE1. Since, we do observe y_1 for all the exposed farmers, the sample average of y_1 from the sub-sample of exposed farmers will consistently estimate ATE1, provided the sample is random (see below). We can decompose ATE as a weighted sum of ATE1 and $E(y_1 \mid w=0)$, the expected adoption impact in the non-exposed subpopulation:

$$ATE = Ey_1 = P(w=1)\times ATE1 + (1-P(w=1))\times E(y_1 \mid w=0) \quad \textbf{(1)}$$

where $P(w=1)$ is the probability of exposure. Hence, once we consistently estimate ATE, ATE1 and the probability of exposure, $P(w=1)$, we can get from (1) the expected "non-exposure" bias $\mathbf{NEB} = P(w=1)\times ATE1 - ATE$; the expected bias from using the sample average adoption rate among the exposed $\mathbf{PSB} = ATE1 - ATE$, and the expected adoption impact in the non-exposed subpopulation $E(y_1 \mid w=0) = \dfrac{ATE - P(w=1)\times ATE1}{P(w=0)}$

As usual, we can obtain the *observed* outcome *y* as function of the potential outcomes y_1 and y_0 and the treatment status variable *w* as:

$$y = wy_1 + (1-w)y_0 = wy_1 \quad \textbf{(2)}$$

Where the second equality follows from the fact that y_0 is always equals to zero for adoption outcomes. Equation (2) shows in particular that the usually computed proportion of adopting farmers, $\dfrac{n_a}{n}$ (where $n_a = \sum_{i=1}^{n} y_i$ is the number of adopting farmers), is a consistent estimator of the joint probability of exposure and adoption $P(wy_1 = 1) = P(w=1, y_1=1)$ and not in general a consistent estimator of the probability of adoption $P(y_1=1)$ even when the sample is random.

Similarly, a logit or probit model $P(y=1|x) = F(x\beta)$, which has observed adoption status *y* as dependant variable and does not condition on *observed* exposure status variable *w*, will not yield consistent estimates of the coefficients of the determinants of *adoption*. At best it will yield consistent estimates of the effects of *x* on the joint probability of exposure *and* adoption, $P(w=1, y_1=1|x)$. But such effect is not informative with regard to the effect of change in *x* on the conditional probability of adoption $P(y_1=1|x)$, which is in principle what a model of determinants of adoption seeks to elicit. Needless to say, the same remarks apply when the intensity of adoption is being modeled through the use of observed land used. We turn next to the estimation of the adoption rate and its determinants based on the observed random vectors $(y_i, w_i, x_i)_{i=1,..,n}$ from a random sample of the population.

The ATE methodology provides the appropriate framework for the consistent estimation of the population adoption rate $E(y_1)$ and that of the determinants of adoption $E(y_1|x)$, which in this framework corresponds to the *conditional* ATE denoted usually as ATE(*x*). Wooldridge (2002, chapter 18) provides a succinct summary of the different estimators available for the consistent estimation of ATE. The ATE estimators are classified under two broad classes based on the assumption they require to be consistent. The first class of estimators is based on the conditional independence assumption. This assumption states that the treatment status *w* is independent of the potential outcomes y_1 and y_0 conditional on an observed set of covariates *x*. The second class of estimators is based on *instrumental variable* methods and assumes the existence of at least one instrument **z** that explains treatment status but is redundant in explaining the outcomes y_1 and y_0, once the effects of the covariates *x* are controlled for. The estimators using the conditional independence assumption are either a pure parametric regression-based

method where the covariates are interacted with treatment status variable, or they are based on a two-stage estimation procedure where the conditional probability of treatment $P(w=1|x)$ ≡ $P(x)$, called the *propensity score*, is estimated in the first stage and ATE and ATE1 are estimated in the second stage by parametric regression-based methods or by non-parametric methods (including the so-called matching methods).

Nonparametric and parametric estimation of adoption rates and their determinants.

The fact that in the adoption context the potential outcome $y_0 = 0$ for both the treated and untreated sub-populations brings several simplifying results compared to the general case. The first simplifying result, which has been mentioned earlier is that the adoption rate among the exposed can be identified and consistently estimated from a random sample of observed adoptions outcomes and exposure status $(y_i, w_i)_{i=1,..,n}$ with no need for additional data or assumptions by $\hat{ATE}1 = \frac{1}{n_e}\sum_{i=1}^{n_e} y_i$, where n_e is the sample number of exposed farmers. In particular, the conditional independence and IV assumptions are not necessary.[4] The second simplifying result leads to a simpler expression of the non-parametric estimate (the so called weighting estimate) of the population adoption rate ATE under the conditional independence assumption:

Proposition 1: If in addition to the conditional independence assumption we assume that the propensity score $p(x) \equiv P(w=1|x)$ satisfies the condition $p(x) > 0$ for all *x*, then the population adoption rate *ATE*= $E(y_1)$ is non-parametrically identified and given by:

$$ATE = E(y_1) = E\left(\frac{y}{p(x)}\right) \quad \textbf{(3)}$$

[4] The proofs of the stated results and propositions are omitted from the paper in order to respect the size limitation. They are available upon request.

Furthermore, *ATE* is consistently estimated from a random sample of observed $(y_i, w_i, x_i)_{i=1,..,n}$ by:

$$\hat{ATE} = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{\hat{p}(x_i)} \quad \textbf{(4)}$$

where $\hat{p}(x)$ is a consistent estimate of the propensity score evaluated at *x*.

In proposition 1, the propensity score $p(x)$ can be estimated by non-parametric methods or by parametric methods such as probit or logit models.[5] The third simplifying result that follows from the fact that $y_0 = 0$ for both the exposed and non-exposed sub-populations concerns the estimation of *ATE*(*x*):

Proposition 2: Under the conditional independence assumption, ATE(*x*) is non-parametrically identified and can be consistently estimated from a random sample of $(y_i, x_i)_{i=1,..,n}$ drawn from the *exposed* sub-population only.

Proposition 2 brings a major difference compared to the general case where observed (y_i, x_i) data from an untreated control or comparison group is required for the identification and consistent estimation of *ATE*(*x*) and *ATE1*(*x*) using parametric or non-parametric methods even when conditional independence is assumed.

3. ATE estimation of NERICA adoption rates and determinants

The data used in our analysis came from a sample of about 1,500 rice farmers in 50 villages in Cote d'Ivoire. The villages were selected among the villages surrounding five WARDA on-farm research sites in the forest and savanna regions. The selection of villages was not random as it purposely included villages where WARDA has been conducting on-farm and Participatory Varietal Selection (PVS) research activities. Indeed, prior to their official release in Cote

[5] The asymptotic distribution of the weighting estimator with a first stage probit estimation can be found in its general form in Lee (2005)

d'Ivoire only in late 2001, the diffusion of the NERICA varieties was mostly through WARDA PVS activities (see Dalton, 2004). But, the sample villages also included neighboring villages where WARDA has not conducted any research activity. Thirty rice farmers were randomly selected in each village for a total sample size of approximately 1,500 farmers.[6] The data collected included in particular the farmer knowledge and cultivation of village varieties and socio-demographic information.

Estimation Results and Discussion

We have used the above results to obtain non-parametric and semi-parametric ATE and ATE1 estimates of the NERICA population adoption rates. The predicted propensity score $\hat{p}(x)$ appearing in equation (4) was estimated in a first stage using a probit model of the determinants of the probability of exposure to the NERICAs, $P(w=1|x)$.[7] We have also used Proposition 2 to estimate a probit model of NERICA adoption $P(y_1=1|x)$ (which we call the ATE probit model) that uses the observations from the NERICA exposed subsample only. Then, using the estimated parameter of the ATE probit model the predicted conditional probability of adoption are calculated for both the exposed and non-exposed sub-samples to obtain the ATE and ATE1 estimates of NERICA population adoption rates (and that of joint exposure and adoption). For comparison purposes, we have also estimated a classic probit model of joint exposure and adoption $P(y=1|x)=P(wy_1=1|x)$ (which we call the naïve probit model) using all the sample observations and the same explanatory variables as the ATE probit model.[8]

NERICA adoption rates and determinants

[6] Because of the nature of the study, we restricted the survey to rice farmers only. Non-rice farmers were randomly replaced whenever present in the first random draw. Full description of the sampling methodology and the nature of the information collected is available upon request (see also Diagne, 2006).

[7] The results and discussion of the first stage probit model estimation of the propensity score is not included in the paper because of the limitation imposed on the size of the paper (see Diagne 2006).

[8] We have used STATA version 9.0 for all estimations and inferences.

The results of the estimation are shown in Table 1 and Table 2. Table 2 shows that for the joint exposure and adoption rate all the three different methods of estimation (ATE semi-parametric, ATE probit and Naïve probit,) give the same point estimates as the direct sample computation estimate (4%), which we know from the above results to be consistent with no additional distributional or functional form assumption (only random sampling is assumed). They also yield approximately the same range for the 95% confidence interval (between 3% and 5%). This suggests that the assumptions underlying the three models are plausible as far the estimation of the joint exposure and adoption rate for the whole population and its determinants is concerned. However, only the ATE semi-parametric and ATE probit model give the same estimate of the joint exposure and adoption rate within the presently NERICA-exposed subpopulation as the direct sample computation estimate (36%), which we also know to be consistent with no additional distributional or functional form assumption. The Naïve probit model estimate of 29% for the joint exposure and adoption rate within the presently NERICA-exposed subpopulation with a 95% confidence interval ranging between 19% and 34% (which does not cover the 36% consistently estimated value) is a sign that the model is suffering from attenuation bias due to its use of the full sample in the estimation without controlling for the exposure variable (Yatchew and Griliches, 1985). It is also a sign that its coefficients estimates are not consistent at least for the presently NERICA-exposed subpopulation.

The downward bias of the Naïve probit model estimate of probability of joint exposure and adoption for the NERICA-exposed subpopulation is also an indication that its coefficients estimates are likely to be biased downward for a model of the determinants of adoption. Indeed, Table 2 shows that the estimated coefficients of the naïve probit model are in general significantly lower than that of the ATE probit adoption model. The only exception to this is the

coefficient for the adoption status in 1999, which is significantly higher for the naïve probit model. But, this is in fact the exception that proves the rule because we know that exposure is a necessary condition for past adoption so that when the estimation use the full sample (i.e. including the observation for the non-exposed), the coefficient for the adoption status in 1999 is likely to be picking up the effects of the omitted exposure variable.[9] One note, however, that the coefficients that are statistically significant for the adoption alone model (adoption in 1999, PVS-hosting village status, number of NERICA varieties known in the village, number of NARS varieties known in the village, past participation in PVS trials, being an upland rice farmer, having a secondary occupation, and living in the forest zone) are also significant for the joint exposure and adoption model. The notable exceptions to this are the coefficients for past participation in PVS trials and that for Forest zone, which are not statistically significant for joint exposure and adoption.

Going back to Table 1, the full population adoption rate (ATE), which inform on the demand of the technology by the target population, is estimated to be 28% by the ATE semi-parametric method and 23% by the ATE Probit model. This means that the NERICA adoption rate in Cote d'Ivoire could have been 28% (resp 23%) in 2000 instead of the actually observed 4% joint exposure and adoption rate, if the whole population were exposed to the NERICAs in 2000 or before. The corresponding estimates of the population adoption *gap* (i.e. the non-exposure bias), are -25% and -19%, respectively. The adoption rate among the presently NERICA exposed subpopulation (ATE1) is estimated to be 36% by both ATE models, while the estimated adoption rates for the non-exposed subpopulation are 27% for the ATE semi-parametric and 21% for the ATE Probit. The estimated implied population positive selection bias (PSB) is 8% for the ATE semi-parametric and 13% for the ATE Probit. The PSB estimates

[9] This is the classical correlated omitted variable problem.

are not, however, statistically significantly different from zero at the 5% significance level. In other words, we cannot reject the hypothesis that the presently NERICA-exposed subpopulation is equally likely to adopt the NERICAs as the general population.

We note that in both the ATE semi-parametric and the ATE probit cases the estimated adoption rates for the non-exposed subpopulation are very close to the full population adoption rate estimates. This is what we expect as only 10% of the population has been exposed to the NERICAs. The fact that the ATE probit based estimates for the adoption rate in the full population and in the non-exposed sub-population appears to be biased downward indicates that the probit parameterization and specification is not entirely satisfactory.

4. Conclusion

We can draw several methodological conclusions from the paper with regard to the estimation of adoption rates and their determinants. First, from the data collection point of view, adoption surveys that do not collect information on the farmer exposure to the technologies are unlikely to lead to reliable estimates of adoption rates and their determinants. Second, when the diffusion of a technology in the population is not complete, estimated adoption rates from direct sample computation and from the classical adoption model are implicitly about joint exposure and adoption and do not inform about adoption per se. Furthermore, the estimated "adoption rates" and determinants are severely biased downward. Third, it is the population adoption rate estimated through the average treatment effect (ATE) estimation framework that provides reliable information on the adoption of a technology in terms of its desirability and potential demand by the target population. Fourth, the difference between the observed joint exposure and adoption rate and the population adoption rate estimated through the ATE framework is the

"adoption gap" that exists solely because of the incomplete diffusion of the technology in the population (i.e. the unrealized potential).

The application of the methodological results of the paper to the estimation of the adoption rates for the New Rice for Africa (NERICA) rice varieties in Cote d'Ivoire show that the ATE methodological approach developed in the paper has significant policy implications with respect to judging the intrinsic merit of a new technology in terms of its potential demand and the decision to invest or not in its wide scale dissemination.[10] In particular, we have found that the NERICA adoption rate in Cote d'Ivoire could have been 28% in 2000 instead of the actually observed 4% joint exposure and adoption rate, if the whole population were exposed to the NERICAs in 2000 or before. This justify investing in the dissemination of the NERICA varieties; considering that the 28% is bound to increase significantly in the future as farmers learn more about the characteristics of the NERICAs and become comfortable with their performances.

[10] The methodology applies equally to the estimation of the demand of a new product not universally known in the population.

Table 1: Estimates of NERICA adoption rates and their 95% confidence intervals[1]

	Sample moments estimates	Naïve Probit Joint exposure and adoption model	ATE semi-parametric estimates			ATE Probit adoption model			
Joint exposure *and* adoption rate (Probability of knowledge *and* adoption of at least one NERICA variety):									
in the full population	0.04 (0.03 0.05)	0.04 (0.03 0.04)	0.04	(0.03	0.05)	0.04	(0.03	0.04)	
within the NERICA-exposed subpopulation[2]	0.36 (0.28 0.43)	0.29 (0.19 0.34)	0.36	(0.28	0.43)	0.36	(0.28	0.41)	
NERICA adoption rate (Probability of adoption of at least one NERICA variety):									
In the full population (ATE)			0.28	(0.17	0.64)	0.23	(0.12	0.36)	
Within the NERICA-exposed subpopulation (ATE1)	0.36 (0.28 0.43)		0.36	(0.28	0.43)	0.36	(0.28	0.41)	
Within the subpopulation not exposed to the NERICAs			0.27	(0.15	0.63)	0.21	(0.09	0.38)	*
Estimated population adoption gap: Expected "non-exposure" bias (NEB)*			-0.25	(-0.60	-0.14)	-0.19	(-0.32	-0.09)	*
Expected population selection bias (PSB)**			0.08	(-0.21	0.19) **	0.13	(-0.06	0.26)	**

[1] 95% bootstrap confidence interval in brackets.

[2] It is clear that for the sample moments these are the same as the adoption rates among the exposed. They are also the same for the ATE-based estimates by Proposition 4 . But for the naïve probit these are not necessarily the same.

* Reject the hypothesis of equality with zero at the 5% significance level but fail reject at the 1% significance level (0.01<p-values <0.05).

** Fail to reject the hypothesis of equality with zero at the 5% significance level (p-values>0.05).

Table 2: Probit models of NERICA adoption and Joint exposure and adoption: Coefficients estimates

	ATE Probit adopion model		Naïve Probit Joint exposure and adoption model	
Adoption of NERICA varieties in 1999	2.42	(5.00)**	3.29	(9.62)**
Village PVS	4.52	(2.71)**	0.62	(2.49)*
Number of NERICA varieties known in the village	0.82	(4.02)**	0.31	(3.34)**
Number of traditional varieties known in the village	-0.02	(0.57)	-0.02	(-0.94)
Number of NARS upland varieties known in the village	-0.56	(2.03)*	-0.15	(-2.51)*
Rice area share in 1999	2.76	(1.60)	0.55	(1.13)
Commercialized crops area share (including rice) in 1999	2.43	(1.69)	0.44	(0.93)
Log of total agricultural cash income in 1999	0.42	(1.55)	0.23	(1.16)
Contact with CIDT/GVC	0.87	(1.11)	0.27	(1.02)
Contact with SATMACI/SODERIZ	1.23	(1.94)	0.06	(0.15)
Participation in Participatory Varietal Selection (PVS) trials before 2000	1.29	(2.31)*	0.30	(1.12)
Practice upland rice cultivation	1.66	(2.35)*	1.25	(3.51)**
Log of farm size (maximum total area cultivated in the past 5 years)	-0.18	(0.32)	0.10	(0.31)
Average Household size in the past 5 years	0.05	(1.17)	0.02	(0.59)
Age	-0.05	(1.90)	-0.01	(-0.47)
Has a secondary occupation	1.13	(2.64)**	0.83	(3.42)**
Years of formal schooling	-0.09	(1.39)	0.00	(0.01)
Woman	-0.31	(0.80)	-0.01	(-0.03)
Ethnie Senoufo			1.15	(3.05)**
Forest zone	-5.69	(2.57)*	0.52	(0.97)
Constant term	-5.14	(2.74)**	-6.67	(-5.82)**
Observations	125		1255	
log-likelihood	-30.46		-59.21	
Pseudo R-2	0.63		0.71	
Wald Chi squared	76.34		247.59	
Df	19		20	

Robust z-statistics in parentheses

* significant at 5%; ** significant at 1%

References

Angrist, J.D., G.W. Imben, and D.B. Rubin (1996), "Identification and Causal Effects Using Instrumental Variables", Journal of The American Statistical Association 91, 444-455.

Besley, Timothy and Anne Case (1993) "Modeling Technology Adoption in Developing Countries" *The American Economic Review*, Vol. 83, No. 2, Papers and Proceedings of the 105 Annual Meeting of the American Economic Association. pp. 396-402.

Blundell, Richard. and M. Costa Dias (2002), "Alternative Approaches to Evaluation in Empirical Microeconomics", Institute of Fiscal Studies, University College London, CEMMAP Working Paper CWP10/02.

Dalton, Timothy. J. (2004). "A household hedonic model of rice traits: economic values from farmers in West Africa." Agricultural Economics 31: 149 - 159.

Diagne, A. 2006. "The diffusion and adoption of NERICA rice varieties in Côte d'Ivoire". *The Developing Economies*, Volume 44:2 .

Dimara, E. and D. Skuras, 2003. "Adoption of agricultural innovations as a two-stage partial observability process". *Agricultural Economics*. 28: 187-196.

Feder, G., R.E. Just and D. Zilberman. 1985. "Adoption of Agricultural Innovations in Developing Countries: A Survey" *Economic Development and Cultural Change* 33 (2): :255-295.

Heckman, J. 1996: "Randomization and Social Program Evaluation As An Instrumental Variable", *Review of Economics and Statistics, 77(2), 336-341*.

Heckman J.J., R. J. Lalonde, and J. A. Smith. 1999. "The Economics and Econometrics of Active Labor Markets Programs" In Handbook of Labor Economics 3B. ed. O.C. Ashenfelter, and D. Card, North-Holland

Lee, Myoung-Jae, 2005. *Micro-Econometrics for Policy, Program, and Treatment Effects* Advanced Texts Series in Econometrics, 248 pp. Oxford University Press

Moffitt, Robert. 1991. "Program Evaluation with Nonexperimental Data." *Evaluation Review*. 15(3): 291-314.

Rosenbaum, P.R., and D. R.Rubin (1983), "The Central Role of the Propensity Score in Observational Studies for Causal Effects," Bometrika 70, 41-55.

Saha, L., L. H. Alan and S. Robert. 1994. "Adoption of emerging technologies under output uncertainty". *American Journal of Agricultural Economics*. 76: 836-846.

Sunding, D. and D. Zilberman. (2001) "The Agricultural Innovation Process: Research and Technology Adoption in a Changing Agricultural Sector." In: Bruce Gardner and Gordon Rausser (eds.), Handbook of Agricultural and Resource Economics. Elsevier Science, Amsterdam: North Holland. pp. 207-261.

Wooldridge, J. 2002, *Econometric Analysis of cross section and Panael data*. The MIT press, Cambridge, Massachusetts, USA.