Technological change in smallholder agriculture: Bridging the adoption gap by understanding its source

ALIOU DIAGNE*

Policy and Impact Assessment Program, The Africa Rice Center (AfricaRice), Cotonou, Benin

This paper examines the informational origin of the low adoption rates of modern agricultural technologies in smallholder agriculture in sub-Saharan Africa. It argues that a large part of these low rates can be explained by the fact that many smallholder farmers are unaware of the existence of these technologies. The paper analyzes the structure of the *adoption gap* resulting from this lack of awareness and presents a methodology for estimating that gap and the truly informative adoption rates and their determinants. This methodology is used to estimate the potential adoption rates and adoption gaps of New Rice for Africa (NERICA) and the determinants of NERICA exposure and adoption in four West African Countries: Côte d'Ivoire, Guinea, Benin and Gambia. The estimated adoption gaps of 21% in Côte d'Ivoire, 41% in Guinea, 28% in Benin and 47% in Gambia suggest that NERICA adoption could be increased significantly.

Keywords: technology diffusion; adoption; adoption gap; selection bias; average treatment effect (ATE); NERICA (New Rice for Africa)

JEL codes: C13; O33; Q12; Q16

Cet article examine l'origine informationnelle du faible taux d'adoption des technologies agricoles modernes par les petits exploitants agricoles de l'Afrique sub-saharienne. Celui-ci affirme qu'une part importante de ces faibles taux peut s'expliquer par le fait que beaucoup de paysans ignorent l'existence de ces technologies. Cet article analyse la structure des écarts d'adoption, causés par cette ignorance, et présente une méthodologie pour mesurer ces écarts, et les taux d'adoption réellement informatifs et leurs déterminants. On utilise cette méthodologie pour évaluer les taux d'adoption potentiels, les éscarts d'adoption du Nouveau Riz pour l'Afrique (NERICA), et pour évaluer les déterminants de l'adoption et de l'exposition de NERICA dans quatre pays de l'Afrique de l'Ouest : la Côte d'Ivoire, la Guinée, le Bénin et la Gambie. Les écarts d'adoption estimés, 21% en Côte d'Ivoire, 41% en Guinée, 28% au Bénin et 47% en Gambie, suggèrent que l'on pourrait augmenter l'adoption du NERICA de manière significative.

Mots-clés : *diffusion de la technologie ; adoption ; écart de'adoption ; biais de selectionn ; effet moyen du traitement (*average treatment effect, ATE*, en anglais) ; NERICA (Nouveau Riz pour l'Afrique)*

Catégories JEL : *C13 ; O33 ; Q12 ; Q16*

* Correspondence: A.Diagne@cgiar.org

1. Introduction

The vast majority of farmers in developing countries are smallholders, with an estimated 85% of them farming less than two hectares (World Bank, 2007). Hence, as emphasized in the 2008 World Development Report, the potential of agriculture to contribute to growth and poverty reduction depends on the productivity of smallholder farmers. And raising that productivity will require a much higher level of adoption of new agricultural practices and technologies than presently observed in the smallholder farming population (De Janvry & Sadoulet, 2002; World Bank, 2007). This paper examines the informational origin of the low adoption rates of modern agricultural technologies frequently observed in smallholder agriculture in sub-Saharan Africa. It argues that a large part of these observed low adoption rates can be explained by a simple fact: that many smallholder farmers are unaware of the existence of the technology. This is especially true when the technology is relatively new.

Before proceeding further, we need to clarify the meaning of the terms 'diffusion' and 'adoption' as used in this paper. In most of the voluminous adoption literature the two are used interchangeably. Often, in papers that make the distinction between the two concepts explicitly or implicitly, adoption of a technology is defined at the individual level to mean its *use* while diffusion is defined at the aggregate population level to mean the *propagation of use* of the technology in the population (Feder et al., 1985; Sunding & Zilberman, 2001), in other words, the extent of adoption in the population.[1] Obviously, a technology must be known to someone before it can be used. But no distinction is generally made in the common use of the two concepts between the mere knowledge or *awareness* of the existence of a technology (without necessarily using it) and its use. Such a distinction is made in this paper. As in Diagne (2006) and Diagne & Demont (2007), the adoption of a technology is defined in this paper to mean its use at the individual level or at the aggregate population level. To be more precise, we will speak of adoption status or adoption intensity at the individual level and adoption rate at the aggregate population level. The term diffusion is used strictly in this paper to mean the extent of awareness of the technology in the population (which does not necessarily imply its use).[2]

If the population's awareness of the existence of the technology is not universal, the diffusion rate as commonly used must be understood as the rate of population awareness *and* adoption, which combines information about two different rates: 1) the rate at which the population is being made aware of the technology, which we call the *diffusion* rate in this paper, and 2) the rate at which the part of the population that is aware of the technology is *using* it, which we call the *adoption rate among the exposed* in this paper.[3] The product of the diffusion rate and the adoption rate among the exposed is the *actual* adoption rate that is consistently estimated by the proportion of adopters from a random sample of the population. As argued below, among all these quantities, only the population adoption rate is in general informative about the intrinsic merit of a technology in terms of the extent of its desirability by the target population. The difference between the population adoption rate and the actual adoption rate

[1] There is often a time or space dimension embodied in the common use of the term 'diffusion'.

[2] The implicit assumption in the common definition and use of the term 'diffusion rate' is that the population's exposure to the technology is universal and only the number of individuals adopting (or dis-adopting) it changes through time. Hence, the significance of 'diffusion rate' in this case is really the adoption rate conditional on universal exposure. This is what we call the population adoption rate (or population *potential* adoption rate) in this paper.

[3] We will use the two terms 'awareness' and 'exposure' interchangeably throughout the paper. However, we use 'exposure' to mean 'awareness of the existence of the technology' and not necessarily to imply any knowledge of the characteristics of the technology.

is what we call the population 'non-exposure' bias, which exists solely because of the incomplete diffusion of the technology in the population. It measures in some sense the *unmet* population demand for the technology and will therefore be called simply the *adoption gap*. Thus the title of the paper.

Although pioneers of adoption studies such as Rogers (1983) and Beale and Bolen (1955, cited in Daberkow & McBride, 2003) have emphasized the critical importance of awareness in the adoption process, most empirical studies of adoption have either ignored the issue or dealt with it inappropriately. In fact, with a few exceptions, empirical adoption studies have so far neglected to collect information on farmers' awareness of the technology being studied. The vast majority of agricultural technology adoption studies do emphasize the critical role that access to information plays in the adoption process (for reviews of the literature see Feder et al., 1985; Sunding & Zilberman, 2001) and empirical models of adoption usually include some information related variables (notably access to extension services) to account for that fact (see for example Adesina & Zinnah, 1993; Adesina & Baidu-Forson, 1995).

However, most of the focus on the role of information in the adoption process has been on information about the characteristics and performance of the technology and the farmer's learning process leading to the acquisition of that information (Feder & Slade, 1984; Wozniak, 1993; Cameron, 1999; Batz et al., 1999). But having information about the characteristics and performance of a technology is conceptually and empirically different from merely being aware of its existence. Furthermore, awareness of the existence of a technology is a sine qua non for its adoption (i.e. use), while, in principle, one can start using a new technology while knowing nothing about its characteristics or performance. It is this fact (i.e. that awareness is a prerequisite for adoption) that makes accounting for awareness fundamental in adoption studies, especially when the technology studied is relatively new. In particular, the usually computed sample adoption rate is uninformative with respect to the expected population adoption rate when only a few farmers are aware of the existence of the new technology and we want to know the extent to which the new technology satisfies the population's demand for new technologies. In fact, as shown in Diagne (2006) and Diagne & Demont (2007), the *observed* sample adoption rate is a consistent estimator of the combined rate of awareness and adoption.

Conflating the awareness and adoption information in the same rate makes it impossible to infer from the observed sample adoption rate the *potential* population adoption rate which, from a policy perspective, is the quantity that tells us about the intrinsic value of the technology to society and the desired policy action. In particular, we cannot know whether, in a particular population, a low observed sample adoption rate is the result of a very low potential adoption rate or just low awareness of the existence of the technology. As pointed out by Diagne (2006), these two possible causes lead to contrasting policy implications: a high potential population adoption rate that is masked by a low level of awareness points to the need to put more effort into extension to make the technology known and available to the larger population. On the other hand, if the potential population adoption rate is low, further extension effort to disseminate the technology may not warrant its cost.

Similarly, if empirical models of the determinants of adoption do not account for the awareness status of farmers, then they are not informative about the factors favoring or constraining adoption, except where awareness of the technology in the population is universal. In other words, one cannot consistently estimate the effects of the factors influencing adoption in such models. Indeed, such models are fundamentally unidentified, meaning essentially that the significations of the quantities they estimate (coefficient estimates and marginal effects) are different from what most think they are. The fundamental difficulty in interpreting the coefficients and marginal effects estimates using the classic adoption model of the determinants of adoption has been pointed out by several authors, including Feder et al. (1985), Besley and Case (1993), Saha et al. (1994) and Dimara and Skuras (2003). In fact, we will show in this paper that these coefficients and marginal effects estimates based on the classic models indeed have different meanings and can be very different from the meaning arrived at by estimating the 'true' adoption function, which correctly and appropriately isolates the effect a factor has on adoption per se from its effect on the awareness status of a farmer. In particular, for the same data and variables the marginal effects estimates from a classical adoption model can be 10 to 100 times smaller than those derived from the correctly specified adoption model.[4] It goes without saying that such a large difference in magnitude and change in statistical significance will in most cases make a qualitative and significant difference to the conclusions one reaches from an adoption study.

The fact that awareness is a necessary condition for adoption also has important implications for how the farmer awareness status information, when available, is accounted for in adoption models. Indeed, an adoption model that does not handle the awareness status variable properly will quickly run into computational difficulties and will not produce results in most cases (i.e. the model estimation will end with an error message in most statistical software).[5] Or, if results are produced (with the aid of a specific functional form that artificially circumvents the problem), chances are that they will be grossly at odds with common sense and the basic facts because of the fundamental unidentifiability of the model. This is the case, for example, when the awareness variable enters additively in the observed adoption function directly or indirectly through a non-linear transformation.[6]

The paper is organized as follows. Section 2 uses the finite population approach and a simple adopter/non-adopter type framework to illustrate and explain the 'non-exposure' bias problem, which is the source of the adoption gap when exposure to the technology is not universal. Section 3 uses the counterfactual outcomes and *average treatment effect* estimation

[4] This empirical finding is understandable as one can show theoretically under some identifying assumptions that the conditional mean 'adoption' function estimated in the classic adoption model is equal to the true population average conditional adoption function (the 'true' population adoption function) multiplied by the probability of being aware of the technology. Hence, for a factor determining adoption alone and not awareness, its marginal effect calculated from the classic 'adoption' model is equal to its marginal effect from the true adoption model multiplied by the conditional probability of awareness, a quantity always between 0 and 1 and usually very small when not many farmers are aware of the technology. For a factor that is determinant of both adoption and awareness, the marginal effect calculated from the classic 'adoption' model will be equal to the same product above plus a second term (made of the marginal effect with respect to awareness multiplied by the 'true' population adoption function). This second term makes the comparison of the two marginal effects theoretically indeterminate. However, in practice the second term will usually be small for most factors in most data.

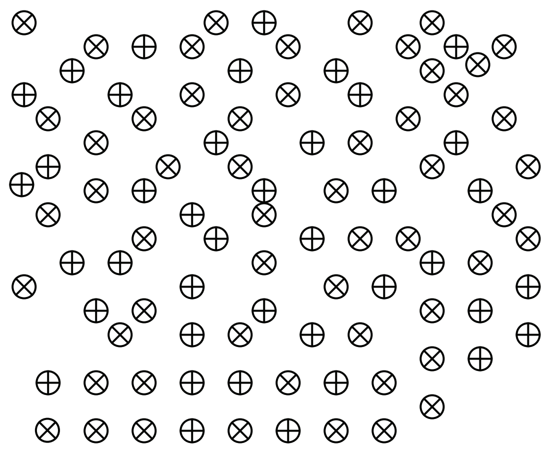
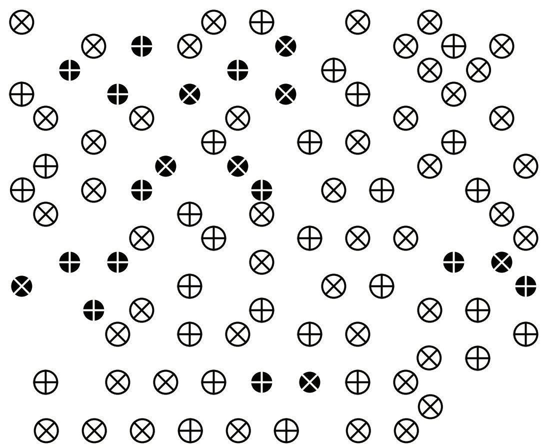
[5] This computational problem is well known in the statistical literature.

[6] That is, the relation between awareness and adoption implies that it cannot be specified as $E(A \mid X = x, W = w) = g(\alpha w + \beta x)$ where A is the observed adoption status variable, W is the individual awareness status variable (equal to 1 if the individual is aware of the technology and 0 otherwise), α and β are parameters and g is a (possibly nonlinear) real valued function. This fact is overlooked by Daberkow and McBride (2003) in their empirical analysis of the influence of awareness on adoption. This alone can explain the 'strange' results of their empirical model which made them conclude that awareness of precision agriculture technologies is not a determinant of their actual adoption. Daberkow and McBride tried to rationalize their findings but the conclusion reached clearly contradicts the fact that awareness is a necessary condition for actual adoption.

framework to show how consistent non-parametric and parametric estimators of population adoption rates and their determinants can be obtained within this framework. Section 4 applies the results of Section 3 to consistently estimate the population adoption rates and determinants of the NERICA (New Rice for Africa) rice varieties in Benin, Côte d'Ivoire, Gambia and Guinea along with estimates of the population 'adoption gap' and selection biases created by the presently limited diffusion of the NERICA varieties. Section 5 concludes the paper with a summary of the major methodological and empirical results of the paper and their policy implications.

2. Anatomy of the source of the adoption gap: A finite population approach

To assess as simply as possible the magnitude of the non-exposure bias in commonly used sample adoption rate estimates, we use a finite population approach and focus on a population of farmers of size N, which can be divided into two groups based on a farmer's adoption attitude toward a given technology: an adopter-type group of farmers who will adopt the technology if exposed to it and an non-adopter-type group who will not (see Figure 1A). We assume that the type of farmer is revealed only through exposure to the technology (see Figure 1B). In other words, one cannot know if a farmer is an adopter type or not until he or she is exposed to the technology. Let N^a be the number of adopter-type farmers and $R^a = \frac{N^a}{N}$ be the proportion of adopter types in the total population. Hence, R^a would be the true population adoption rate when exposure is complete in the population (i.e. when the entire population has been exposed to the technology).

A Population before exposure ⊕ adopter type ⊗ non-adopter type		Total population size: N = **100** Adopter type sub-population size: N^a = **40** Non-adopter type sub-population size: N_0^a =**60** Expected population adoption rate: $R^a = \frac{N^a}{N}$ = **40%**
B Population after partial exposure ● exposed type ○ non-exposed type		Exposed subpopulation size: N_e = **20** Number of adopters among the exposed: N_e^a = **12** Non-exposed subpopulation size: $N - N_e$ =**80** Number of adopter type among the non-exposed: N_0^a =**28** Population exposure *and* adoption rate: R_1^a =**12%** Population exposure rate: R_e = **20%** Adoption rate among the exposed: R_e^a = **60%** Adoption rate among the non-exposed: R_{0e}^a =**35%** Non-exposure bias: **NEB= -28%** Population selection bias: **PSB=+20%**

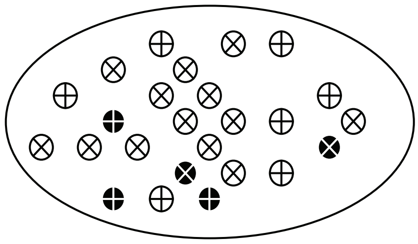
C Random sample from the partially exposed population		Sample size: n =**25** Sample number of exposed: n_e=**5** Sample number of (*revealed*) adopters: n_a=**3** Sample adoption rate: $\frac{n_a}{n}$ = **12%** Sample exposure rate: $\frac{n_e}{n}$ =**20%** Sample adoption rate among the exposed: $\frac{n_a}{n_e}$ = **60%**

Figure 1: Population adoption and joint exposure *and* adoption rates under partial exposure to a technology and positive population selection bias

Now suppose that the population is only *partially* exposed to the technology and let N_e be the size of the exposed subpopulation and $R_e = \frac{N_e}{N}$ be the corresponding exposure rate. Let also N_e^a be the number of adopters in the exposed subpopulation and N_0^a the number of adopter-type farmers in the non-exposed subpopulation with $R_1^a = \frac{N_e^a}{N}$ and $R_0^a = \frac{N_0^a}{N}$ being the corresponding respective proportions in the total population and $R_e^a = \frac{N_e^a}{N_e}$ the proportion of adopters in the exposed subpopulation. Thus, the group of adopter-type farmers is further partitioned by the partial exposure into two subgroups: one subgroup with farmers whose types are revealed and another whose types are still unknown (see Figure 1B). The group of non-adopters is partitioned similarly.

It is important to note that the observable quantities in the above definitions are the total population size N, the size of the exposed subpopulation N_e and the number of adopters in the exposed subpopulation N_e^a.[7] We cannot observe the total number of adopter types in the total population and in the non-exposed subpopulation N^a and N_0^a, respectively. So we cannot compute the true population adoption rate R^a. We can only directly compute the proportion of *revealed* adopters in the population R_1^a, the exposure rate R_e and the proportion of adopters in the exposed subpopulation R_e^a. However, since $N^a = N_e^a + N_0^a$, the knowledge of either N^a or N_0^a allows the computation of the other. The same applies for R^a and R_0^a because $R^a = R_1^a + R_0^a$. In the example illustrated in Figure 1, N = 100, N_e = 20 and

[7] This observability assumes, of course, the feasibility of surveying the whole population.

$N_e^a = 12$ are observable. But $N^a = 40$ and $N_0^a = 60$ are not observable. Thus the true population adoption rate R^a = 40% and the proportion of non-exposed adopter-types in the population R_0^a = 28% cannot be directly known.

With a random sample of farmers, the three observable population parameters (R_1^a, R_e^a and R_e) are consistently estimated by their respective sample analogues (see Figure 1C).[8] In particular, the usually computed sample adoption rate (i.e. the proportion of sample farmers who have adopted) consistently estimates R_1^a but not the true population adoption rate R^a as is commonly believed.

Given the definitions and notations above, we have:

$$R^a = R_1^a + R_0^a = \frac{N_e}{N} \times \frac{N_e^a}{N_e} + \left(1 - \frac{N_e}{N}\right) \times \frac{N_0^a}{(N - N_e)} = R_e R_e^a + (1 - R_e) R_{0e}^a \qquad (1)$$

where R_e^a and R_{0e}^a are the adoption and would-be adoption rates in the exposed and non-exposed subpopulations, respectively.

The right-hand side of the last equality of equation (1) shows that the true population adoption rate is the weighted average of R_e^a and R_{0e}^a, in the exposed and non-exposed subpopulations' adoption rates, respectively, with the weights given by the respective subpopulation shares.[9] But, more importantly, equation (1) shows that taking the sample analogue of R_1^a, the proportion of revealed adopters in a sample, as estimate of adoption rate, generally leads to underestimation of the true population adoption rate R^a. As a measure of population adoption rate R_1^a is incomplete in the sense that it does not take into account the would-be adopters whose types are not revealed. In the example illustrated in Figure 1, we have R_1^a = 12%, which understates the true population adoption rate (40%) by 28%.

We can see from equation (1) that the expected adoption gap or 'non-exposure' bias, defined as **GAP** ≡ $R_1^a - R^a = -(1 - R_e)R_{0e}^a$, is strictly negative and diminishing with increasing exposure rate. This shows that the incomplete population adoption rate R_1^a always understates the true population adoption rate R^a, unless either the exposure rate is equal to 1 or the would-be adoption rate in the non-exposed subpopulation R_{0e}^a is zero.

We can also obtain from equation (1) the expected bias resulting from using the sample analogue of R_e^a, the adoption rate in the exposed subpopulation, as estimate of the true population adoption rate R^a. This expected bias, which is caused by the fact that exposure to the technology is not random in the population, is given by **PSB** ≡ $R_e^a - R^a$ =

[8] The zero-mean sampling error is ignored in the example for clarity.

[9] It should be noted that normally both R_e^a and R_{0e}^a depend on the exposure rate R_e. But we are omitting showing this dependence to simplify the notation.

$(1-R_e)(R_e^a - R_{0e}^a)$. Because the population selection bias **PSB** can be either positive or negative depending on the relative magnitude of the two in subpopulation adoption rates R_e^a and R_{0e}^a, R_e^a can overestimate or underestimate R^a. Overestimation occurs when the adoption rate in the exposed subpopulation is greater than that of the non-exposed one. Otherwise we have underestimation. The population selection bias vanishes only when there is complete exposure or when the exposed and non-exposed subpopulation adoption rates are equal.

In the example illustrated in Figure 1, where the true population adoption rate is 40%, the relatively low population exposure rate of 20% leads to a population adoption gap of -28% and a positive population selection bias of +20%. The dependence of the population adoption gap and selection biases on the population exposure rate is illustrated in Figure 2 under positive (A), negative (B) and zero (C) population selection biases, respectively.[10]

Figure 2A: The positive population selection bias case: the subpopulation most likely to adopt is exposed first

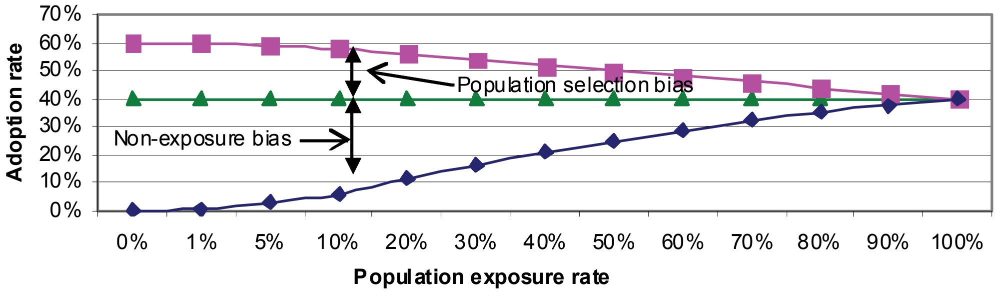
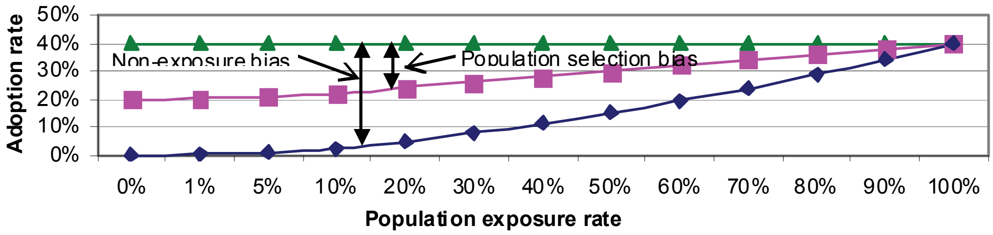


Figure 2B: The negative population selection bias case: the subpopulation least likely to adopt is exposed first



[10] In Figure 2 it is assumed that the deviation of the adoption rate in the exposed subpopulation from the true population adoption rate as a result of a population selection bias is a linear function of the exposure rate. That is, $\frac{\Delta R_e^a}{R^a} \equiv \frac{R_1^a - R^a}{R^a} = \alpha(1-R_e)$, where α is the constant population selection bias parameter, with a positive value indicating a positive population selection bias and a negative value the opposite. With this linear functional form assumption we have $R_e^a = (1+\alpha(1-R_e))R^a$ and $R_1^a = R_e R_e^a = (1+\alpha(1-R_e))R_e R^a$ (α=0.5 in Figure 2A and α=-0.5 in Figure 2B).

Figure 2C: The zero population selection bias case: all subpopulation members are equally likely to be exposed

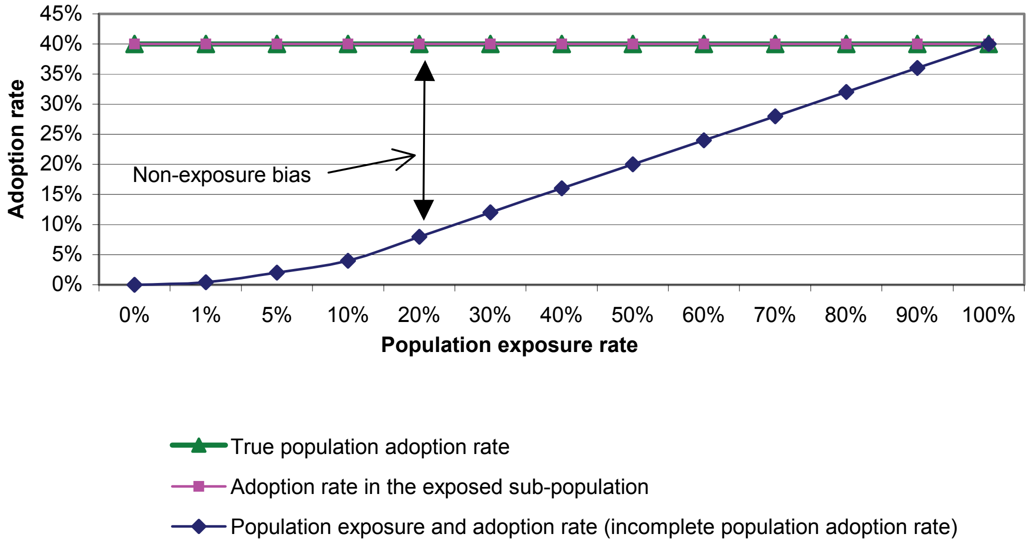


Figure 2: Population adoption rates and non-exposure and selection biases as function of exposure rate

We can see from equation (1) and the preceding discussion that the sample proportion of revealed adopters is in fact an estimate of the population joint exposure *and* adoption rate (JEA). Indeed, **JEA** $\equiv R_1^a$ is exactly the proportion of farmers in the total population who are exposed to the technology *and* who have adopted it. Therefore, such sample adoption rate estimate embodies two types of information: about the diffusion of a technology and about its adoption. We could, however, argue that the question that interests us in an adoption study is the extent to which farmers *like* a given technology and not the extent to which they *know* about it. Indeed, it is the answer to the question of how much a technology is liked that provides feedback to researchers about the suitability of their research product for meeting the needs of the targeted population. The answer to the question 'How well known is a technology?' is most useful for assessing the performance of extension systems or methods. If we conflate the two different types of information, then the sample proportion of revealed adopters provides little information about the potential population adoption rate when exposure is low. In fact, as pointed out by Manski (2005) in a more general context, it is clear from the right-hand side of equation (1) that all that can be learned about the true population adoption rate R^a from the empirical evidence without any assumption is that it lies between the observed joint exposure and adoption rate $R_1^a = R_e R_e^a$ (when the unobserved R_{0e}^a is zero) and the value $R_1^a + 1 - R_e$ (when the unobserved R_{0e}^a is 1). This interval is, however, usually too wide to be informative.

We will see in the next section that within the treatment effect estimation framework the true population adoption rate R^a is consistently estimated by the so-called *average treatment effect* (ATE), whereas the *average treatment effect on the treated (*ATT) consistently estimates R_e^a, the adoption rate in the exposed subpopulation.

3. Average treatment effect (ATE) estimation of population adoption rates and their determinants

As shown by Diagne (2006) and Diagne and Demont (2007), the ATE methodology enables the identification and consistent estimation of the population mean adoption outcome $E(y_1)$ which, with exposure as the treatment and y_1 being the potential outcome under exposure, corresponds to the average treatment effect (ATE).[11] The average treatment effect on the treated (ATT) and the average treatment effect on the untreated (ATU), which correspond here to the respective mean potential adoption outcomes for the exposed and non-exposed subpopulations, are also identified by the conditional expected value $E(y_1 \mid w=1)$ and $E(y_1 \mid w=0)$, respectively; where w is a binary variable indicating the observed status of exposure to the technology (with $w=1$ if the farmer is exposed and $w=0$ otherwise). In this context, the observed adoption outcome y is given by $y = wy_1$, with the population mean joint exposure and adoption parameter (JEA), the population adoption gap (GAP) and the population selection bias (PSB) given respectively by

JEA $= E(y) = E(wy_1)$, $GAP = E(y) - E(y_1)$, $PSB = E(y_1 \mid w=1) - E(y_1)$.[12]

One approach to the identification of ATE is based on the so-called conditional independence assumption (Wooldridge, 2002: Ch. 18; Imbens & Wooldridge, 2009) which states that the treatment status w is independent of the potential outcomes y_1 conditional on the observed set of covariates x that determine exposure (w). The ATE parameters identified through the conditional independence assumption can be estimated from a random sample of observed $(y_i, w_i, x_i)_{i=1,..,n}$ in two different ways:[13] 1) using a weighting estimator and 2) using an estimator based on a parametric regression procedure.

The inverse probability weighting (IPW) estimator of ATE

The weighting estimator is based on a two-stage estimation procedure where the probability of treatment conditional on a covariate vector *z*, $P(w=1 \mid z) \equiv P(z)$, called the propensity score, is estimated in the first stage and ATE, ATT and ATU are estimated in the second stage using the following probability weighting estimators, which are special cases of the general weighting estimators of ATE, ATT and ATU when $y_0 = 0$ (Diagne & Demont, 2007):

[11] This follows from the fact that with exposure as treatment the potential outcome under non-exposure is equal to zero because without exposure there is no adoption.
[12] We should note that when the adoption outcome variable is a binary variable taking the values 0 and 1, as is the case in our empirical analysis, then the expected values corresponding to the various population mean adoption outcomes reduce to probability quantities that correspond to measures of population adoption *rates* (i.e. proportions of adopting farmers in the population).
[13] One can also use matching-based estimators (see for example Imbens & Wooldridge, 2009).

$$A\hat{T}E = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{\hat{p}(z_i)} \tag{2}$$

$$A\hat{T}T = \frac{1}{n_e}\sum_{i=1}^{n_e} y_i \tag{3}$$

$$A\hat{T}U = \frac{1}{n-n_e}\sum_{i=1}^{n}\frac{(1-\hat{p}(z_i))}{\hat{p}(z_i)}y_i \tag{4}$$

where $\hat{p}(z)$ is a consistent estimate of the propensity score evaluated at z and $n_e = \sum_{i=1}^{n} w_i$ is the sample number of exposed farmers.[14]

Parametric estimation of ATE

The parametric estimation procedure of ATE is based on the following equation that identifies the average treatment effect conditional on a vector of covariates x under the conditional independence (CI) assumption (see Diagne & Demont, 2007):

$$ATE(x) = \mathrm{E}(y_1 \mid x) = \mathrm{E}(y \mid x, w = 1) \tag{5}$$

The parametric estimation proceeds by first specifying a parametric model for the conditional expectation in the right-hand side of equation (4), which involves the observed variables *y*, *x* and *w*:

$$E(y \mid x, w = 1) = g(x, \beta) \tag{6}$$

where g is a known (possibly nonlinear) function of the vector of covariates x and the unknown parameter vector β which is to be estimated using standard Least Squares (LS) or Maximum Likelihood Estimation (MLE) procedures using the observations (y_i, x_i) from the

[14] The propensity score $p(z)$ can be consistently estimated using non-parametric methods or parametric methods such as probit or logit models (see Imbens & Wooldridge, 2009). We note that the weighting estimator for ATT is simply the proportion of adopters in the exposed subsample and does not depend on the estimated propensity score $\hat{p}(z_i)$. Also implicit in the weighting estimators is the requirement that $0 < \hat{p}(z_i) < 1$ and $0 < n_e < n$.

subsample of exposed farmers only, with y as the dependent variable and x the vector of explanatory variables. With an estimated parameter $\hat{\beta}$, the predicted values $g(x_i, \hat{\beta})$ are computed for all the observations i in the sample (including the observations in the non-exposed subsample) and ATE, ATT and ATU are estimated by taking the average of the predicted $g(x_i, \hat{\beta})$ $i=1,..,n$ across the full sample (for ATE) and respective subsamples (for ATT and ATU):

$$A\hat{T}E = \frac{1}{n}\sum_{i=1}^{n} g(x_i, \hat{\beta}) \qquad (7)$$

$$A\hat{T}T = \frac{1}{n_e}\sum_{i=1}^{n} w_i g(x_i, \hat{\beta}) \qquad (8)$$

$$A\hat{T}U = \frac{1}{n-n_e}\sum_{i=1}^{n} (1-w_i) g(x_i, \hat{\beta}) \qquad (9)$$

The effects of the determinants of adoption as measured by the K marginal effects of the K-dimensional vector of covariates x at a given point $\bar{x}$ are estimated as:

$$\frac{\partial E(y_1 \mid \bar{x})}{\partial x_k} = \frac{\partial g(\bar{x}, \hat{\beta})}{\partial x_k} \quad k = 1,..,K \qquad (10)$$

where x_k is the k^{th} component of x.

In our empirical analysis below, we have estimated the ATE, ATT, ATU, the population adoption gap ($G\hat{A}P = J\hat{E}A - A\hat{T}E$)[15] and the population selection bias ($P\hat{S}B = AT\hat{E}1 - A\hat{T}E$) parameters using both the inverse probability score weighting (IPSW) estimators (equations 1, 2 and 3) and the parametric regression based estimators (equations 4, 5 and 6). The propensity score $\hat{P}(z)$ appearing in the IPSW estimators is estimated using a probit model of the determinants of exposure: $P(z) = \Phi(z\gamma)$ where Φ is the standard normal cumulative distribution with density function $\phi(t) = (1/\sqrt{2\pi})\exp(-t^2/2)$, z the observed vector of covariates determining exposure to the technology and γ is the parameter vector being estimated. This estimation of the determinants of exposure is important for its own sake as it can provide valuable information about the factors influencing farmers' exposure to a new technology. These factors, which are mostly related to the diffusion of information, can very well be different from those influencing the adoption of the technology once exposed to it. For the

[15] Note that as discussed earlier the joint exposure and adoption parameter (JEA) is consistently estimated by the sample average of the *observed* adoption outcome values: $J\hat{E}A = \frac{1}{n}\sum_{i=1}^{n} y_i$.

parametric regression based estimators, since *y* is a binary variable in our empirical analysis, the equation 5 above is effectively a parametric probabilistic model as we have discussed earlier. That is, we have $E(y \mid x, w=1) = P(y=1 \mid x, w=1)$ with, assuming a probit model, $g(x,\beta) = \Phi(x\beta)$. Thus in this particular case the parametric estimation of ATE reduces to a standard probit estimation restricted to the exposed subsample. The marginal effects in equation (9) are also estimated using this ATE parametric model.[16] For comparison purposes, we have also estimated a 'classic' probit adoption model (which, as discussed above, is in fact a model of the determinants of joint exposure and adoption): $P(y=1 \mid x') = \Phi(x'\theta)$ where $x' = (z, x)$ is the vector of covariates determining both exposure (*w*) and adoption (y_1) and θ is the parameter vector to be estimated.[17] All the estimations were done in Stata using the Stata add-on *adoption* command developed by Diagne (2007) to automate the estimation of ATE adoption models and related statistical inference procedures.[18] The asymptotic distributions of $\hat{ATE}$, $\hat{ATT}$ and $\hat{ATU}$ are given in Lee (2005:67–9) for the general case where $y_0 \neq 0$ and $p(z)$ is estimated through a probit model.

4. ATE estimation of NERICA adoption rates and their determinants

The NERICA (New Rice for Africa) rice varieties, developed by AfricaRice in 1990s, are the result of interspecific crosses between *Oryza sativa* rice species from Asia and the locally adapted and multiple-stress resistant *Oryza glaberrima* African rice species. From 1996, NERICA varieties were introduced in many African countries through participatory varietal selection (PVS) trials and were then disseminated by farmers through their informal channels. In this section we estimate NERICA diffusion rates and their actual and population adoption rates and gaps in Côte d'Ivoire, Guinea, Gambia and Benin where they were introduced, starting in 1996 (Côte d'Ivoire and Guinea) and 1998 (Gambia and Benin). The determinants of NERICA diffusion and adoption in these four countries are also estimated.

Sampling and data

The data used in the paper are collected from samples of about 1,500 rice farmers in 50 villages in Côte d'Ivoire in 2000, 1,467 rice farmers in 79 villages in Guinea in 2001, 360 rice farmers in 24 villages in Benin in 2004 and 600 rice farmers in 70 villages in Gambia in 2006. A multi-stage stratified random sampling method was used to select the sample rice farmers in all four countries, with the last two stages consisting of selecting the sample villages and farmers located in all the regions where NERICA has been introduced. The selection of sample villages was, however, not entirely random as it purposely included

[16] Note that the marginal effects of the determinants of adoption (i.e. the effects of the marginal changes in the vector of covariate *x*) cannot be estimated from the IPSW based estimators.

[17] We should note that usually the two vectors *z* and *x* have common elements so that the dimension of the vector x' is usually less than the sum of the dimensions of its two components. It is clear that by not having in the vector x' determinants of *w* not in *x* will most likely result in the non-identification of 'classic' adoption models. However, in practical estimation terms the main difference between the ATE parametric adoption model and the 'classic' adoption model lies in the fact that the latter uses *all* the sample observations while the former uses the observations from the exposed subsample only.

[18] The *adoption* command is a Stata add-on command that works like standard Stata regression commands. It uses various Stata standard estimation commands internally to implement the estimation procedures described above and, depending on the option chosen, provide IPSW or parametric regression based estimates of ATE, ATT, ATU, JEA, GAP and PSB.

villages where AfricaRice has been conducting on-farm and PVS research activities. In selecting the sample villages, a list of all villages where NERICA seed was introduced (called NERICA villages) was drawn up first. The sample NERICA villages were then randomly selected from that list. Then, for each sample NERICA village, a list of neighboring villages within five to ten kilometers of where NERICA was not introduced (called non-NERICA villages) was constituted and one or two sample villages were randomly selected from that list. Thus, in Côte d'Ivoire, 25 NERICA villages and 25 non-NERICA villages were selected in the forest and savanna regions. In Benin, 12 NERICA villages and 12 non-NERICA villages were selected in the central region. In Gambia, 35 NERICA villages and 35 non-NERICA villages were selected in all the four agricultural regions of the country. In Guinea, the villages were selected among four agro-ecological zones where NERICA dissemination activities were being conducted. In each zone a further stratification was done into two types of prefectures: NERICA prefecture (where NERICA varieties had been introduced) and non-NERICA prefectures (where NERICA varieties had not yet been introduced). For each NERICA village selected, three to four non-NERICA villages in the surroundings of that NERICA village were selected. In all four countries, farmers were selected entirely randomly from the population of rice farmers in the sample villages, with the sample size varying across countries: 30 per village in Côte d'Ivoire, 15 per village in Benin and 20 per village in Guinea. In Gambia, in some villages 5 farmers were selected, and in others 10.[19]

In each country, the data was collected at both village and farmer levels through a structured questionnaire. At the village level, the data collected included the rice varieties known in the village (modern and traditional) and village infrastructures and community variables. At the farmer level, the data included the rice varieties known and cultivated by the farmer and other socio-demographic data. Prior to administering the farmer level questionnaire, a list of the known varieties in the village was constructed from the village level survey and each sample farmer was asked about his or her own knowledge and cultivation of the varieties known in his or her village.

Demographic and socioeconomic characteristics of sample farmers

Table 1 reports selected descriptive statistics of the sample farmers in the four countries disaggregated by their adoption status. Common variables have been chosen[20] for the purpose of comparison and brevity. The table shows that non-adopters and adopters of NERICA in each country have approximately the same average age (except in Côte d'Ivoire, where adopters seem to be older than non-adopters, but the difference is not statistically significantly different from zero). The mean household size is higher in Gambia (16) than in the other countries (6 in Benin, 7 in Côte d'Ivoire and 10 in Guinea). The differences in household size between adopters and non-adopters are not statistically different from zero, however, except in the case of Guinea. The same pattern is also evident for the education level of the household's head, with adopters reporting significantly more years of formal education than non-adopters, except for Guinea. There are no significant differences between adopters and non-adopters across the four countries in professional training or in type of experience in rice farming.

Table 1: Demographic and socioeconomic characteristics of adopters and non-adopters

[19] 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.
[20] Results including the non-common variables are available upon request.

Characteristics	Benin		Côte d'Ivoire		Gambia		Guinea	
	Adopters	Non-adopters	Adopters	Non-adopters	Adopters	Non-adopters	Adopters	Non-adopters
Age	42.98	42.72	43.77	40.90	45.19	45.13	48.72	48.05
Household size	7.2	5.7***	7.7	6.9	16.5	15.9	10.9	9.9**
Years of schooling			2.98	2.23	4.20	3.82	4.40	4.80
Percentage of women	58	61	43	34	92	94	3	7
Percentage of men	42	39	57	66	08	6	97	93
Contact with extension	70	60	9.4	9.8	14	14	62	40
Literacy (reading and writing)	8	5			33	37	1.4	1.7
NERICA village	74	44	68	52	68	38	58	38
Non-NERICA village	26	56	32	48	32	62	42	62

Note: In Guinea, Côte d'Ivoire and Benin, the figures represent the percentage of female-headed households. The information on the years of schooling was not collected in Benin, only literacy information (yes or no) was collected.

The results in Table 1 show that women are the large majority of rice growers in Gambia (more than 90%) but they constitute a very small minority in Guinea (less than 5%). The proportion of female adopters in the sample is lower than female non-adopters, except in Côte d'Ivoire. The proportion of sample farmers with access to extension services is relatively high in Benin (more than 60%) and Guinea (more than 40%) compared to the other two countries. There are also more NERICA adopters with access to extension in these two countries compared to non-adopters, whereas in Gambia and Côte d'Ivoire the proportions of farmers with access to extension service are about the same for adopters and non-adopters. As can be expected, in all four countries the proportion of NERICA adopters is higher in the NERICA villages than in the non-NERICA villages.

Results of the ATE estimation of NERICA adoption rates and gaps

The results of the estimation of the different NERICA diffusion and adoption rates and gaps are presented in Table 2. The NERICA diffusion (i.e. exposure) rates are estimated to be 9% for Côte d'Ivoire in 2000, 39% for Guinea in 2001, 26% for Benin in 2004 and 57% for Gambia in 2006. Abstracting from country differences in NERICA dissemination efforts, we can see from these estimates a steady progress of NERICA diffusion from 2000 to 2006. Table 2 also shows that the estimation of the population joint exposure and adoption rates (JEA) using the two different ATE methods of estimation (IPSW estimator and ATE probit) yields the same estimates as the directly computed sample adoption rates for all the four countries, and the estimates are statistically significant at the 1% level with broadly the same ranges for the respective 95% confidence intervals. These joint exposure and adoption rates are 4% for Côte d'Ivoire, 20% for Guinea, 19% for Benin and 40% for Gambia. As demonstrated above, because of the relatively low diffusion of the NERICA varieties in all the four countries, these joint exposure and adoption rates estimates significantly understate the population adoption rate (i.e. the adoption rate that would be obtained if the whole population were exposed to the NERICA varieties).

Table 2: Estimates of NERICA adoption rates in the four countries and their 95% confidence intervals in parenthesis

Parameters	Sample moment estimates	Inverse propensity score weighting (IPW) estimator of ATE	ATE probit adoption model
Exposure rate			
Benin	0.26 (0.07)		
Côte-d'Ivoire	0.09 (0.08)		
Gambia	0.57 (0.12)		

Guinea	0.39 (0.08)		
Actual adoption rate (adoption and exposure)			
Benin	0.18 (0.06)	0.19 (0.14 0.24) ***	0.19 (0.14 0.24) ***
Côte-d'Ivoire	0.04 (0.005)	0.04 (0.03 0.05)	0.04 (0.03 0.05)
Gambia	0.40 (0.03)	0.40 (0.36 0.43) ***	0.40 (0.36 0.43) ***
Guinea	0.20 (0.04)	0.20 (0.18 0.22)	0.20 (0.18 0.22)
Potential adoption rate (ATE)			
Benin		0.45 (0.33 0.56)***	0.47 (0.37 0.57)***
Côte-d'Ivoire		0.22 (0.10 0.35)***	0.24 (0.15 0.34)***
Gambia		0.85(0.75 0.95)***	0.87 (0.83 0.91) ***
Guinea		0.61 (0.50 0.71)***	0.61 (0.56 0.65)***
Adoption rate among exposed (ATT)			
Benin		0.52 (0.35 0.69)***	0.52 (0.43 0.61)***
Côte-d'Ivoire		0.37 (0.27 0.47)***	0.38 (0.30 0.45)***
Gambia		0.86 (0.70 1.02)***	0.86 (0.81 0.90)***
Guinea		0.55 (0.47 0.64)***	0.55 (0.51 0.59)***
Adoption rate among non-exposed (ATU)			
Benin		0.41 (0.29 0.52)***	0.44 (0.32 0.56)***
Côte-d'Ivoire		0.21 (0.07 0.34)***	0.23 (0.13 0.33)***
Gambia		0.84 (0.75 0.92)***	0.88 (0.84 0.92)***
Guinea		0.64 (0.50 0.78)***	0.64 (0.59 0.70)***
Adoption gap			
Benin		-0.26 (-0.33 -0.19)***	-0.28 (-0.36 -0.20)***
Côte-d'Ivoire		-0.19 (-0.31 -0.07)***	-0.21 (-0.30 - 0.12)***
Gambia		-0.45 (-0.50 -0.41)***	-0.47 (-0.50 -0.45)***
Guinea		-0.41 (-0.50 -0.32)***	-0.41 (-0.45 - 0.38)***
Population selection bias (PSB)			
Benin		0.07 (-0.02 0.17)	0.05 (-0.01 0.11)**
Côte-d'Ivoire		0.15 (0.03 0.26) **	0.13 (0.04 0.22) ***
Gambia		-0.01(-0.08 0.10)	-0.01 (-0.02 -0.00)**
Guinea		-0.05 (-0.14 0.04)	-0.06 (-0.09 -0.03) ***

As shown in Table 2, the adoption rates in the NERICA-exposed subpopulation (ATT) are estimated to be 52% for Benin, 37% for Côte d'Ivoire, 86% for Gambia and 55% for Guinea, with approximately the same respective ranges for the 95% confident intervals for the two methods (IPSW and ATE probit). The estimates are all statistically significant at the 1% level of significance. As explained above, these adoption rates among the NERICA-exposed subpopulation are likely to overstate the NERICA (potential) population adoption rates because of positive selection bias.

The NERICA population adoption rates (ATE), which offer information about the target population's demand for NERICA, are estimated to be 45%, 22%, 85% and 61% for Benin, Côte d'Ivoire, Gambia and Guinea, respectively, by the IPSW method and 47%, 24%, 87% and 61%, respectively, by the ATE probit model. The estimates are all statistically significantly different from zero at the 1% level of confidence. It can be seen that for each country the ATE probit method shows in general adoption rates estimates that are 2% higher than those of the IPSW method, except for the Guinea case, where the two estimates are the same. We note also that the probability of adopting at least one NERICA variety is highest in Gambia and lowest in Côte d'Ivoire.

The corresponding estimates of the NERICA population adoption *gap* (i.e. non-exposure bias) as given by the IPSW and ATE probit methods are respectively -26% and -28% in Benin, -19% and -21% in Côte d'Ivoire, -45% and -47% in Gambia and -41% in Guinea, with all the estimates statistically significant at the 1% level. These adoption gap estimates imply that there is still potential for increasing NERICA adoption rates significantly in all four countries. It should be emphasized that this adoption gap is solely due to the lack of awareness of the existence of NERICA. However, the size of the adoption gap depends on the same factors that determine the exposure and population adoption rates, the effects of which are estimated below. Hence, by appropriately changing the values of these determinants through some policy instruments, one can increase actual adoption through a simultaneous narrowing of the adoption gap and an increase in the population adoption rate.

The (potential) adoption rates in the subpopulation not exposed to the NERICA varieties (ATU) are estimated by the IPSW and ATE probit methods to be 41% and 44% in Benin, 21% and 23% in Côte d'Ivoire, 84% and 88% in Gambia and 64% in Guinea. The estimated implied population selection bias (PSB) is 7% and 5% in Benin, 15% and 13% in Côte d'Ivoire, -1% in Gambia and -5% and -6% in Guinea for the IPSW and ATE probit methods, respectively. The PSB estimates are all significantly different from zero at least at the 5% level for all countries in the case of the ATE probit model. This implies that the probability of adoption for a farmer belonging to the subpopulation of exposed farmers is significantly different from the probability of adoption for any other farmer randomly selected in the general population. The negative PSB for Guinea and Gambia indicates that the farmers exposed to the NERICA varieties are significantly less likely to adopt at least one NERICA variety than any farmer randomly selected from the population (which indicates mis-targeting of the NERICA dissemination activities).

Determinants of NERICA exposure and adoption

In this section we present and discuss the results of the estimation of the probit model of the determinants of exposure to (i.e. awareness of) the NERICA varieties and that of the determinants of NERICA adoption in the population from the parametric ATE probit model. The results of the estimation of the classic probit adoption model (which is in fact a model of the determinants of joint exposure and adoption, as shown above) are also presented for comparison purposes.

Table 3: Exposure probit model coefficient estimates (coef.) and marginal effects (dy/dx) in the four countries

	Benin		Côte-d'Ivoire		Gambia		Guinea	
Variables	**Coef**	**dy/dx**	**Coef.**	**dy/dx**	**Coef.**	**dy/dx**	**Coef.**	**dy/dx**
NERICA village	0.92***	0.32***	0.92***	0.43***	0.73***	0.30***	0.39***	0.11***
Village contact with the existing extension service (ANADER)			-0.2					
Village contact with NGO			-0.21					
Village past contact with the past extension service (no longer existing)			1.57***	0.075***				
Number of NERICA varieties known in the village	0.163		0.51***	0.023***			0.462***	0.130***
Number of traditional varieties known in the village			0.043***	0.002***			-0.002	-0.001
Total number of varieties from the National Agricultural Institute known in the village			-0.13*	-0.006*	0.054	0.019	0.012	0.003
Past participation in PVS trials			0.1	0.210*				
Practice upland rice cultivation			0.89***	0.09				
Log of total land size	0.367		0.16	-0.03				
Household size	0.097**	0.031**	0.01				0.011	0.003
Being born in the same village			0.24					
Age	0.015	0.005	0.003	0			0.0065	0.002
Having a secondary activity			0.280*	0.09				
Years of formal schooling	-0.193	-0.062	0.039*	-0.01				
Being female	0.395	0.127	0.29	0			-0.467*	-0.122**
Being from Bete ethnic group			-0.99***	0.249*				
Being from Senoufo ethnic group			1.2	0.340**				
Being in Forest zone			0.91					
Farmer contact with the past extension service (no longer existing)				0.04				
Average household size for the past five years				0.01				
Literacy in traditional language	0.403	0.135						
Within a farmer association	-0.106	-0.034						
Receiving training on rice	0.08	0.026						
Western Region					0.588***	0.210***		
Contact with the National Agricultural Research Institute (NARI)					0.562*	0.196*		
Number of years resident in village							-0.002	-0.001
Middle Guinea							-0.046	-0.013
Upper Guinea							0.612***	0.190***
Forest Guinea							-1.189***	-0.287***
Experience in upland rice farming							0.178	0.05

Experience in lowland rice farming				-0.400***	-0.108***
Village contact with the NGO Sasakawa Global 2000				0.173	0.05
Farmer contact with the past extension service				0.352***	0.10**
Constant	-2.331*	-6.19***	-0.748***	-1.59***	
Number of observations	268	1261	600	1467	
Pseudo R2	0.13	0.37	0.11	0.235	
Chi square	46.97	296.42	89.65	448.19	
Degrees of freedom	9	20	4	15	
Log of likelihood	-151.35	-257.09	-369.3	-731.48	
Akaike information criterion	322.7	556.17	748.59	1494.9	

Determinants of NERICA exposure

Table 3 presents the results of the exposure probit model and the marginal effects of the determinants of the probability of being exposed to the NERICA varieties in Côte d'Ivoire, Guinea, Benin and Gambia, respectively. The results show that across all four countries, living in a NERICA village (where the PVS activities were conducted) is the most important determinant of exposure to the NERICA varieties. Access to extension services is also an important determinant of exposure for Côte d'Ivoire and Guinea. The results also show that rice farmers living in a village with relatively larger number of NERICA and traditional varieties are significantly more likely to be exposed to NERICA and that farmers who practice the upland rice farming system are more likely to be exposed to the NERICA varieties, which is understandable given the fact that NERICA is an upland variety. It is notable that in Guinea women are less likely than men to be exposed to the NERICA varieties. This suggests that the dissemination activities may have been biased against women and that more targeting of women should be done. In Gambia, the fact that living in the Western Region was found to be an important determinant of exposure is not surprising since the first PVS activities were located in this region of the country. And the fact that the headquarters of the National Agricultural Research Institute (NARI), are located in the same region also explains the high probability of exposure to NERICA in the region.

Table 4: Marginal effect estimates obtained from ATE and classic probit models for the four

Variables	**Benin**		**Côte-d'Ivoire**		**Gambia**		**Guinea**	
	ATE	**Classic**	**ATE**	**Classic**	**ATE**	**Classic**	**ATE**	**Classic**
NERICA village	0.012	0.153***	0.427***	0.057***		0.234***	-0.024	0.057*
Village contact with NGO				-0.01				0.029
Village contact with the past extension service (no longer existing)			0.075***	0.217***				
Number of NERICA varieties known in the village			0.023***	0.077***				0.075***
Number of traditional varieties known in the village			0.002***	0	-0.017***	-0.008***		0

Total number of varieties from the National Agricultural Institute known in the village			-0.006*	-0.032***				
Past participation to PVS trials			0.210*	0.033*				
Practice upland rice cultivation			0.09	0.037***			-0.032	0.018
Log of total land size	0.122*	0.044*	-0.03	0				
Household size	0.007	0.017*		0	-0.002	-0.001	0.002	0.002
Being born in the same village				0.01				
Age	0.064**	0.025*	0	0			-0.001	0.001
Having a secondary activity			0.09	0.026*				
Being female	-0.094	0.027	0	0.01			-0.264**	-0.138***
Being from Bete ethnic group			0.249*	-0.01				
Being from Senoufo ethnic group			0.340**	0.15				
Being in Forest zone				0.04				
Farmer past contact with the past extension service (no longer existing)			0.04	0.01				
Literacy (reading and writing)		0.064						
Attended a meeting where research trials results are discussed		-0.03						
Receiving training on rice		0.012						
Age squared	-0.001**	-0.000*						
Log of total land size								
Number of rice varieties known	0.054	0.019						
Number of modern varieties in the village						0.016		0.002
Contact with the National Agricultural Research Institute (NARI)						0.172*		
Western Region					-0.192**	0.066		
Experience in lowland					-0.043	-0.021	-0.071	-0.090***
Farmer past contact with main extension service (Gambia)					0.093**	0.045**		
North Bank Region					-0.217**	-0.103**		
Number of years resident in village							0.002	0.001
Middle Guinea							-0.187*	-0.074*
Upper Guinea							-0.255***	-0.002
Forest Guinea							0.213*	-0.126***
Farmer past contact with main extension service (Guinea)							0.169***	0.125***
Total number of varieties from the National Agricultural Institute							0.087***	0.033***

(IRAG) known by farmer		
Total number of traditional varieties known by farmer	-0.028***	-0.011***

Determinants of NERICA adoption

Table 4 presents the results of the estimated coefficients and marginal effects of the ATE probit adoption model and the classical probit adoption model for Côte d'Ivoire, Guinea, Benin and Gambia. As explained above, one should keep in mind when interpreting the results that the classic 'adoption' model is really a model of joint exposure and adoption. The results in the tables show in general marked differences between the two models in the magnitudes and statistical significances of the coefficients and the marginal effects. The differences are particularly striking for the case of Côte d'Ivoire, where the marginal effects of the ATE probit model are up to 100 times larger in absolute values than those of the classic 'adoption' model.

We can draw some important conclusions about the major determinants of NERICA adoption across countries from the results in Table 4. First, one can see from the table the importance of where the farmer lives (for example, in a NERICA village in Côte d'Ivoire, in the forest area of Guinea, or in the Western or North Bank Region of Gambia) in affecting positively and significantly the probability of NERICA adoption in all the four countries, except Benin. These places of residence are often the villages or the regions where NERICA varieties have been introduced. However, it must be emphasized that the effect of the farmer's village and place of residence is on the probability of adoption per se and independent of the positive effect that the introduction of NERICA in those places has on the actual adoption (i.e. joint exposure and adoption) through increasing the probability of farmers being aware of them. These location specific effects reveal the importance of PVS activities in these villages for increasing farmer knowledge of the characteristics of the NERICA varieties. Indeed, farmers living in these areas, even if they are not participating in the PVS trials, can more easily visit the NERICA trials by themselves or learn about the NERICA varieties by discussing them with PVS participants. This explanation is reinforced by the finding that direct participation in PVS trials has a significant and positive effect on NERICA adoption in Côte d'Ivoire. These location specific effects also reveal the suitability of the NERICA varieties to those identified regions compared with the others; information which is very useful for targeting purposes in dissemination activities.

Second, the results also show the importance of access to extension services in determining NERICA adoption. The positive contribution of such access is consistent with prior expectations and the general findings in the literatures (Feder et al., 1985; Sunding & Zilberman, 2001). This is also consistent with the role of extension as an important source of information about the characteristics and performance of varieties for farmers and the importance such information plays in the five stages of the adoption process proposed by Rogers (1983): (a) knowledge, (b) persuasion, (c) decision, (d) implementation and (e) confirmation. However, the results of the study show that the number of farmers who have access to extension advice remains relatively low in these countries, which suggests that there is scope for increasing the cultivation of NERICA by intensifying extension efforts. This is particularly important for Guinea and Gambia, where agricultural extension workers have had a significant impact in persuading farmers to adopt the NERICA varieties in addition to creating awareness of them among farmers.

Third, the negative correlation between being female and NERICA adoption found in Guinea points to some possible gender biases in the way the NERICA varieties disseminated in Guinea were selected and introduced in that country. First, it is well known that the various NERICA lines tested in the PVS and on farm trials in Guinea differed in some key characteristics that are of importance to women (ease of threshing, for example). It may well be that the NERICA lines that were ultimately selected for release and seed multiplication were the ones that satisfied mostly the varietal characteristics that male Guinean rice farmers preferred (high potential yield, for example). This finding therefore calls into question the suitability of the NERICA varieties disseminated in Guinea for the particular needs of women rice producers. Second, and related to this point, Guinean extension workers, most of them men, may have focused their extension efforts on male farmers to the point that not much information about the differing characteristics and performance of NERICA varieties promoted in Guinea reached the women. This finding is consistent with an observation made by Lo (2000) that despite their role as the backbone of the farm household's food production and consumption in the Sahel, women have limited access to critical resources, technology inputs and support services such as credit and extension due to cultural, traditional and sociological factors. The World Bank (1995) also notes that rural women in the Sahel are not frequently reached by extension services and are rarely members of cooperatives, which often distribute government subsidized inputs to small farmers. Still, consistent with this notion, Kinkingninhoun-Mêdagbé et al. (2008), in their analysis of the impact of gender discrimination on productivity and technical efficiency in Benin, observe that female rice farmers in Benin are particularly discriminated against when it comes to access to production resources, which has a negative effect on their productivity and income. Hence there is need to take a closer look at the gender composition of extension services in Guinea and the way male extension agents work in rice farming communities.

5. Conclusion

This paper has shown the importance of appropriately controlling for exposure and selection bias when assessing the adoption rates of a new technology and its determinants. It has argued that a major source of the commonly observed low level of adoption of modern technologies in smallholder farming in sub-Saharan Africa is smallholder farmers' lack of awareness of the existence of the technologies. The structure of the adoption gap resulting from this lack of awareness was analyzed in the paper and a methodology for estimating that gap and truly informative adoption rates and their determinants based on the ATE framework was presented and discussed. This methodology was then used to estimate the NERICA population potential adoption rates and gaps and the determinants of NERICA exposure and adoption in four West African Countries: Côte d'Ivoire, Guinea, Benin and Gambia.

From a methodological point of view, four major conclusions about the way adoption studies are conducted can be drawn from the analysis in the paper. First, from a data collection point of view, adoption surveys must collect information about farmer awareness of the existence of technologies. Otherwise they 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 classic adoption model are implicitly about joint exposure and adoption and do not inform about adoption per se. Third, it is the population adoption rate estimated through the ATE estimation framework that provides reliable information about 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 results from the lack of awareness of the existence of the technology, which we argue is the main cause of the observed low adoption rates of modern agricultural technologies in smallholder agriculture in Africa.

The results of the analyses of the determinants of NERICA exposure and adoption in Côte d'Ivoire, Guinea, Benin and Gambia show that, had the whole rice farming population of these four countries been exposed to the NERICA varieties at the time of the surveys, their adoption rates could have been much higher. The implied estimated adoption gaps suggest a potential for increasing NERICA adoption significantly in these four countries.

The results of the analysis of the determinants of NERICA adoption illustrate the importance of controlling appropriately for awareness in adoption models. Three main empirical findings emerge from that analysis. First, simply conducting PVS trials in a village promotes the adoption of NERICA beyond the subpopulation participating in the trials, most likely because it enables non-participating farmers to learn about the characteristics and performance of varieties from the participating farmers. This beneficial effect on adoption is in addition to the positive effect it has on actual adoption through increased awareness of the existence of the varieties. Second, the importance of farmer access to extension services in promoting NERICA adoption was also a very important finding of the study. Like the PVS trials, access to extension services enables farmers to learn the characteristics and relative performance of varieties after they are made aware of their existence.

One major policy implication stemming from these empirical findings with regard to NERICA adoption in the four countries is the need to invest further in the dissemination of NERICA by enabling extension services and NGOs to reach more rice farmers and provide them with relevant information about these varieties so as not only to bridge the existing adoption gap but also to push further the potential adoption rate of NERICA. Another major policy implication of the findings is the need to promote and invest in PVS as an effective tool for developing new varieties with high chances of adoption by farmers. Furthermore, the finding that PVS encourages adoption beyond the participating farming population suggests a potentially cost-effective strategy for scaling up PVS, consisting of conducting it in more communities with fewer participating farmers.

This paper has dealt only with the gap between actual and potential adoption created by lack of awareness. However, a larger adoption gap is similarly created by lack of access to a technology as this is also a necessary condition for adoption, while awareness is a necessary condition for access. However, the same methodology can be used to estimate the gap due to lack of access when such information is collected. The methodology can be applied in the area of consumer demand and product market research to obtain a reliable estimate of the potential demand for a new product (not known by all consumers) and its determinants.

Acknowledgements

This paper draws on joint collaborative work with the National Agricultural Systems of the member countries of the Africa Rice Center (AfricaRice). The author is grateful to the late Marie Jose Sogbossi for help with this work; his collaborators in the National Agricultural Systems, Florent Kinkingninhoun-Mêdagbé and Gnonna Soul-kifouly Midingoyi for research assistance; Perpetue Kouame and Simon Ncho for help with data processing; his AfricaRice colleagues for helpful discussion; and the guests editor and participants at the 2009 AERC conference in Mombasa for their input. Financial support from the European Union, the International Fund for Agriculture, the Government of Japan, the United Kingdom and the African Development Bank is gratefully acknowledged.

References

Adesina, A & Baidu-Forson, J, 1995. Farmers' perceptions and adoption of new agricultural technology: Evidence from analysis in Burkina Faso and Guinea, West Africa. Agricultural Economics 13(1), 1–9.

Adesina, A & Zinnah, M, 1993. Technology characteristics, farmer perceptions, and adoption decisions: A Tobit model application in Sierra Leone. Agricultural Economics 9(4), 297–311.

Batz, J, Peters, K & Janssen, W, 1999. The influence of technology characteristics on the rate and speed of adoption. Agricultural Economics 21(2), 121–30.

Beale, G & Bolen, J, 1955. How farm people accept new ideas. Cooperative Extension Report No. 15, Iowa State University, Ames, Iowa.

Besley, T & Case, A, 1993. Modeling technology adoption in developing countries. The American Economic Review 83(2), 396–402.

Cameron, L, 1999. The importance of learning in the adoption of high-yielding variety seeds. American Journal of Agricultural Economics 81(1), 83–94.

Daberkow, SG & McBride, WD, 2003. Farm and operators characteristics affecting the awareness and adoption of precision agricultural technologies in USA. Precision Agriculture 4(2), 163–77.

De Janvry, A & Sadoulet, E, 2002. World poverty and the role of agricultural technology: Direct and indirect effects. Journal of Development Studies 38(4), 1–26.

Diagne, A, 2006. The diffusion and adoption of NERICA rice varieties in Côte d'Ivoire. The Developing Economies 44(2), 208–31.

Diagne, A, 2007. Stata model for estimating ATE based adoption. The Africa Rice Center, Cotonu, Benin.

Diagne, A & Demont, M, 2007. Taking a new look at empirical models of adoption: Average treatment effect estimation of adoption rate and its determinants. Agricultural Economics 37(2–3), 201–10.

Dimara, E & Skuras, D, 2003. Adoption of agricultural innovations as a two-stage partial observability process. Agricultural Economics 28(3), 187–96.

Feder, G & Slade, R, 1984. The acquisition of information and the adoption of new technology. American Journal of Agricultural Economics 66(3), 312–20.

Feder, G, Just, RE & Zilberman, D, 1985. Adoption of agricultural innovations in developing countries: A survey. Economic Development and Cultural Change 33(2), 255–95.

Imbens, G & Wooldridge, JM, 2009. Recent developments in the econometrics of program evaluation. Journal of Economic Literature 47(1), 5–86.

Kinkingninhoun-Mêdagbé, FM, Diagne, A, Simtowe, F, Agboh-Noameshie, AR & Adégbola, PY, 2008. Gender inequality and its impact on income, productivity and technical efficiency: Evidence from Benin. Agriculture and Human Values 27(1), 57–69.

Lee, M-J, 2005. Micro-Econometrics for Policy, Program and Treatment Effects (Advanced Texts in Econometrics). Oxford University Press, New York.

Lo, HM, 2000. Integrating conflict management into community forestry: Experience from West and Central Africa. Institut des Sciences de l'Environnement, UCAD (University Cheikh Anta Diop of Dakar), Dakar, Senegal.

Manski, CF, 2005. Social Choice with Partial Knowledge of Treatment Response. Princeton University Press, Princeton and Oxford.

Rogers, E, 1983. Diffusion of Innovations. Third edition. The Free Press, New York.

Saha, L, Alan, LH & Robert, S, 1994. Adoption of emerging technologies under output uncertainty. American Journal of Agricultural Economics 76(4), 836–46.

Sunding, D & Zilberman, D, 2001. The agricultural innovation process: Research and technology adoption in a changing agricultural sector. In Gardner, B & Rausser, G (Eds), Handbook of Agricultural and Resource Economics. Elsevier Science, Amsterdam, pp. 207–62.

Wooldridge, J, 2002. Econometric Analysis of Cross Section and Panel Data. The MIT press, Cambridge, MA.

World Bank, 1995. Rural Women in the Sahel and their Access to Agricultural Extension: Overview of Five Country Studies. Sector Study, The World Bank, Washington, DC.

World Bank, 2007. World Development Report 2008. Agriculture for Development. Washington, DC.

Wozniak, GD, 1993. Joint information acquisition and new technology adoption: Late versus early adoption. Review of Economics and Statistics 75(3), 438–45.