Technology Adoption in Poorly Specified Environments

Pilar Useche ♣

Department of Food and Resource Economics and Center for Latin American Studies

University of Florida

useche@ufl.edu

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, July 29-August 1, 2007

Abstract

This article extends the characteristics-based choice framework of technology adoption to account for decisions taken by boundedly-rational individuals in environments where traits are not fully observed. It is applied to an agricultural setting and introduces the concept of ambiguity in the agricultural technology adoption literature by relaxing strict informational and cognition related assumptions that are implied by traditional Bayesian analysis. The main results confirm that ambiguity increases as local conditions become less homogeneous and as computational ability, own experience and nearby adoption rates decrease. Measurement biases associated with full rationality assumptions are found to increase when decision makers have low computational ability, low experience and when their farming conditions differ widely from average adopter ones. A complementary empirical paper (Useche 2006) finds that models assuming low confidence in observed data, ambiguity and pessimistic expectations about traits predict sample shares better than models which assume that farmers do not face ambiguity or are optimistic about the traits of new varieties.

♣ The author would like to thank Brad Barham, William Brock, Jean-Paul Chavas, Jeremy Foltz and William Provencher at University of Wisconsin-Madison for invaluable comments and suggestions.

Technology Adoption in Poorly Specified Environments

1. Introduction

The analysis of agricultural technology adoption decisions has typically relied on models where agents are assumed to be fully rational.[1] These models depict decision-makers similar to the ones described by Debreu (1959), "who have a complete plan of action made now for the whole future", over completely specified (physically, temporally, and spatially) commodities. In these models, all relevant information necessary to make an adoption (or non-adoption) decision is available to the decision maker at no cost, he has perfect cognition and forms expectations taking all available information into account, without making any systematic errors when predicting the future.

In the 1950's Herbert Simon (1955, 1959) called the attention of the economics profession to the need to introduce limits on human knowledge and computational ability into their models of rationality. However, mainstream technology adoption theory in economics has not yet been significantly affected by this call. The likely reason why this has not happened is that theories based on full rationality, even though unrealistic in many settings, provide a complete picture of behavior under uncertainty -- commodities can be contingent on future events (McFadden 1999, p. 6). Since uncertainty is a quintessential feature of innovation processes --and even more so in agricultural settings— researchers seeking to explain adoption of agricultural innovations find this framework attractive because it allows for the resolution of uncertainty through complex patterns of experience and learning.

However, it is precisely in agricultural settings with rapid technological innovation where the rational framework might be less realistic and where learning might not completely resolve uncertainty. First of all, many agricultural technologies have traits that interact with the ecosystem and with other technologies, changing across space and over time. Thus, agricultural technological alternatives might not be completely specfied in Debreu's sense, at a moment in time. Second, information in agricultural settings tends to be less perfect and more costly than in

[1] See for example Feder and O'Mara 1981, 1982; Besley and Case 1993; 1994; Foster and Rosenzweig 1995; Cameron 1999; Isik and Khanna 2003; Munshi 2004

typical consumer demand scenarios (Experimentation and information gathering are more costly than with standard consumer goods precisely because of the interactions between the traits of agricultural technologies with other factors that often are unobservable). Third, important determinants of technological outcomes or the outcomes themselves might be unobservable or only partially observable (both to farmers and to the researcher).[2] And fourth, lower education levels in rural areas and lower availability of computational devices might translate into lower ability to process information and perform complicated cognitive operations. For example, Munshi (2004) notes that available information sometimes appears to be persistently ignored by individuals or communities.

Furthermore, the rational framework is more consistent with a laboratory type situation where the decision-maker perceives the environmental conditions to be identical or replicable than with a changing environment like the agricultural one. In the standard rational setting, the uncertainty faced by decision makers is regarded as arising from known *iid* stochasticity in key outcomes rather than arising out of systematic unobservability or biases in the observability of these outcomes. That is, though expectations may turn out incorrect, the deviations will not depart systematically from the expected values observed by the researcher on a random sample of the population. This helps resolve uncertainty and, together with the assumption that key outcome distributions are stationary over time, helps agents learn about true distribution parameters and completely resolve uncertainty. After many trials, the decision-maker "will naturally become confident that an observed empirical frequency (of e.g. positive results) is close to a "true" frequency (of positive outcomes), that can be used for forecasting future trials … That is, he will eventually become confident enough to view the data as an i.i.d. process" (Eppstein and Schneider (2006), p.3).

In this article, I seek to develop tools to understand the choices made by agents with imperfect information about choice traits, with limited cognitive ability for information storing and processing, and who have to make decisions in an uncertain, vaguely specified or poorly understood environment. Contrary to the common view that models of microeconomic behavior

[2] Unobservability or partial observability of outcome distributions occur when it is impossible to infer the distribution parameters even from an infinite number of observations (Manski 1995). As an example, late adopters of an agricultural technology might observe the outcomes obtained by early adopters, but might never be able to infer the true distribution of outcomes that is representative for a farm randomly drawn of the population. If the farming conditions of early adopters are fundamentally different from the conditions of late adopters (and the differences in farming conditions are determined by heterogeneous unobserved factors), regardless of how many trials are analyzed neither the researcher nor the farmer can know the full distribution (or infer it).

with bounded rationality are less rich and deep than those achieved by standard models assuming full rationality, [3] my view is that substantive rationality is actually a model constraint in this type of scenario. I apply my research to the choice of adoption of genetically modified crop varieties by farmers. In an effort to understand how expectations about technological traits (such as profitability or management intensity) determine agricultural technological choices, most authors have portrayed farmers as subjects learning through a *Bayesian* belief-updating mechanism. However, the *rationality constraints* on adoption behavior actually implied in the application of this mechanism are left tacit and have not been explicitly analyzed. In this work I specify a generalized framework of the technology adoption problem under uncertainty and analyze these constraints.

One main aspect of constrained rationality analyzed in this study relates to the question of what agents can know about technological traits in complicated environments --where important relevant factors are unobserved, which affect outcomes and information disseminated about them, and where these factors vary over time in ways that are not well understood by decision makers--. Standard rational frameworks have modeled decision-makers as being overconfident and ambitious about what they can know (Epstein and Schneider (2006). This is especially the case in situations where agents have no reason to believe that trial conditions can be replicable, where several potential outcomes seem possible (Shackle 1955), or when agents are not able to or are not confident enough about being able to measure uncertainty (Knight 1921). In all of these cases, decision-makers might make choices under *ambiguity*. That is, when their choice problem is not completely or uniquely defined (Ellsberg 1961; Gilboa and Schmeidler 1989; Epstein and Schneider 2003; Epstein and Wang 2004). My view is that the lack of consideration of the type of environment –simple or complicated-- in which innovations are to be adopted has not only lead to misspecification of the problem but to confounding the role of information and cognitive ability in the adoption process. Often by assuming implicitly that the cognitive ability of decision makers resembles the computational ability of the researcher, or on the other extreme, by simplifying the nature of the problem so much that the analysis of these factors becomes irrelvant.

Extensive research about a subject tightly linked to the choice-under-ambiguity framework has been conducted by Manski (2005, 2004, 2003a, 2003b, 1995). This research

[3] Rubinstein (1998) provides a full review and supports this statement.

emphasizes the fact that exact knowledge of some outcomes (or their probability distribution) cannot be derived even from an infinite number of observations. Thus, the specification of a unique distribution of expectations about those outcomes necessarily involves non-refutable assumptions (if applied to the identification problem faced by the analyst) or a lot of subjective judgement (from the point of view of the farmer). Manski's motivation of an ambiguous scenario, however, is not necessarily related to a complex, changing environment or to subjects who are unable to measure uncertainty (or who do not use numerical probabilities to represent it), but more to the inherent unobservability of existing outcome probability distributions.

Whichever motivation one might choose as a starting point, both lines of research point to the idea that decision-makers in non-experimental scenarios might not be able to exactly evaluate their expected utility from new technologies. This implies that farmers consider a *range* of expected utilities, rather than a specific value. I incorporate this feature in the model and use a choice criterion that was first proposed by Hurwicz (1951). This criterion is consistent with the spirit of boundedly rational agents who use simple decision rules when faced with ambiguity. It differs from standard utility maximization, in the model below, for farmers who are uncertain about their expected utility. It also includes the standard, single valued expected utility maximization, as a special case, for farmers who are confident about their calculations and their data.

Although models of boundedly rational decision-makers have emphasized the limited (or costly) character of information and the less than perfect cognition of agents, they generally assume that these agents are still confident about their information and estimation and/or take decisions in observable or perfectly specified environments. Thus, they can be considered as corresponding to the special case of single-valued expected utility models mentioned above. An example of an approach that considers these aspects, and in fact represents the exception to the rule in the context of technology adoption modeling, is the work by Ellison and Fudenberg (1993). It represents an important departure from the standard Bayesian/rational framework and highlights important weaknesses related to the adoption predictions of the latter.

In this article I propose an individual behavior technology adoption model for boundedly rational agents who also face ambiguity in their utility assessments. I also identify measurement biases in standard adoption models and propose alternative estimation strategies for both multi-valued and single-valued expected utility cases. After specifying a general form of the model, I illustrate how it relates to a standard Bayesian approach and show how looking at the problem through the Bayesian lens brings the researcher back to the full rationality problem. I argue that the major restrictions are introduced when it comes to the assumptions about rationality and belief updating that help operationalize the model. I justify and propose an alternative approach that permits the researcher to make operational the model in a way that seems more sensible for a complicated setting like agriculture.

The present approach proposes that in complicated environments farmers might have difficulties inferring full distributions of traits; thus, they do not form expectations about the statistical deviation of traits and rather try to locate their expected value through rules of thumb. These rules reflect that, in building their expectations about new technologies, farmers care primarily about the qualitative nature of observed information in terms of representativeness (Kahneman and Tversky (1974)) or closeness to their *location*. In the spirit of bounded rationality, decision-makers are featured as being myopic in that they do not aim to maximize more than one-period-ahead expected utility.[4]

Key findings are, first, that farmers' ambiguity, with respect to the decision of whether to adopt a new variety or not, decreases as local conditions are more homogeneous and as computational ability, own experience and popularity of the variety increase. Second, that an ambiguity-averse farmer faced with a decision among very similar varieties might choose based on his familiarity with them or similar varieties. Third, that measurement biases associated with full rationality assumptions increase when decision-makers have low computational/cognitive ability, low experience with a variety or closely related varieties and when their farming conditions differ widely from the average adopter ones. Fourth, standard data exchangeability assumptions have lead to lack of attention to the information sorting and processing undertaken by farmers. Fifth, farmers sorting of information separates the population in heterogeneous groups that differ in their information structures and defines different adoption determinants.

[4] My interest here is not to understand the motion of the problem, but to focus on the additional source of uncertainty introduced by the unobservability of part of the distribution of outcomes. Allowing for a forward-looking agent would not change the essence of the problem, but would either introduce an additional assumption regarding the cognitive ability of the decision-maker (e.g. he would have to compute the expectation over time of traits that are perfectly observed at time t), or add a different source of uncertainty.

2. Conceptual Framework

This article introduces the concept of ambiguity in the agricultural technology adoption literature and links it to boundedly rational agents. It seeks to relax the strict informational and cognition related assumptions of the full rationality framework of choice under uncertainty that are implied by traditional Bayesian analysis. Once one steps outside of those strict informational assumptions, however, a wide ocean of alternatives regarding expectation-beliefs and choice rules are possible. A main contribution to anchoring these alternatives in the terrain of what is humanly possible --as opposed to what is just "possible"- is Shafir (2003), who compiles the main writings about "Preference, Belief and Similarity" by Amos Tversky. In their work regarding judgement under uncertainty Tversky and Kahneman (1971; 1973; 1974; 1982) strongly emphasize that individuals make choices based on data of limited validity, processing it according to heuristic rules that assess likelihood based on similarity and ignore aspects like the relationship between sample size and sample variability.

The study of choice under ambiguity is still a relatively new research area in economics --See working papers by L. Hansen, T. Sargent, L. Epstein and C. Manski—. It derives from two main lines of research that are concerned with unobservable properties of relevant variables in economic analysis. The first line has its origin in the study of the uncertainty faced by individuals making decisions and the second, which mostly analyses identification problems faced by econometricians (partial identification), originated as a study of inference with missing outcome data. Although the concepts of *ambiguity* and *partial identification* differ in their origins, these two lines of research are starting to converge in analyses where both the decision-maker that is being studied faces an uncertainty or unobservability problem and the researcher faces a problem of missing data and identification. However, the fact that both concepts are often equated reflects a commonly blurry line between the decision-problem of the agents under study and the statistical problems faced by the analyst. This problem arises because of the need of the analyst to identify some variable of interest, which leads him to make statistical assumptions that result in assumptions about agents' behavior and cognition. Generally, these imply strong restrictions regarding the scenarios under which the models will provide reasonable predictions.

Gilboa and Schmeidler (1989), Schmeidler (1989) and Epstein and Schneider (2003; 2006) have tried to axiomatize behavior of agents in scenarios where uncertainty might not be exactly measured. These studies are a response to the wide evidence that has been accumulated regarding the inconsistency of the Von-Neumann-Morgenstern expected utility hypothesis in several decision-making scenarios. Other studies have seen this as a problem related to the amount of relevant information that an agent has in the moment that he makes a choice. When agents have too little information they might not even build prior beliefs about a choice attribute (Ellsberg 1961) or might hold multiple beliefs (Gilboa and Schmeidler 1989), in this case, about the traits of a technology. In all of these cases, decision-makers have to make choices under *ambiguity*.

Aspects that contribute to the type of uncertainty that determines ambiguity are limited (or costly) information and less than perfect cognition. These aspects have been analyzed by the bounded rationality literature, but have not necessarily been linked to complicated environments or identification issues (Rubinstein 1989). Ellison and Fudenberg (1993) (E&F onward) considered farmers as boundedly rational agents in two different dimensions. In the context of limited information, they specify an agent's observability range (window width), which is thought of as an informational constraint that impedes agents to observe outcomes at faraway locations, or alternatively, is " the result of players' prior belief that faraway locations are sufficiently different and experiences there are not relevant to their own decisions." (P.614). On the perspective that decision-makers have limited cognition, they specify an information updating mechanism --based on popularity and inertia -- which is much simpler and requires less information than the standard Bayesian one.

Although their's is a population model – for which this work has been criticized because it makes difficult the direct comparison with individual behavior standard rational-actor models - the essential informational, cognitive and behavioral features associated with representative agents are an important departure from the standard Bayesian/rational framework which highlights two main weaknesses related to the adoption predictions of these models. First, in E&F's model, agents reevaluate their choices, such that –in the technology adoption setting- they might disadopt a technology. And second, they allow for enough heterogeneity of agents such that, even under full information, they will not all make the same choice.

In scenarios where there is a range of possible outcomes or where, at least, this is the expectation, choices are not likely to be the result of conventional utility maximization. Manski (2004) has put forward reasonable alternative decision rules in ambiguous settings. One is the *maxmin* rule, proposed by Wald(1945). A second one, proposed by Hurwicz(1951) consists of the maximization of a weighted average of the minimum and maximum values of the objective function that are feasible for each alternative. However, unlike the latter, the maxmin has been made into an axiom by Gilboa and Schmeidler (1989) and shown to correspond to preferences that account for aversion against risk and ambiguity. They showed that these preferences can be numerically represented by a *robust* utility functional (see also Follmer and Schied 2004; Section 2.5). They interpret the non-exact knowledge of a probability distribution as resulting in a set of prior models, which possibly describe the probabilities of future scenarios. This type of preferences have only recently started to be used in applied analysis, in particular, and increasingly, in analysis of optimal investment decisions (Talay and Zheng 2002, Quenez 2004, Schied 2004, 2005; Burgert and Ruschendorf 2005, Muller 2005, and Hernandez-Hernandez and Schied 2006).

In order to focus on the factors determining ambiguity and how these interact with rules of thumb by boundedly rational agents, I do not consider strategic interactions among agents, which have been subject of study by several authors in the technology adoption literature (Bandiera and Rasul; Udry and Conley). However, social effects might have some bearing in determining the way in which farmers sort out information and strategic delays in adoption could be seen as the result of the ambiguity faced by farmers. For example, Bandiera and Rasul (2005) find that in Mozambique adoption choices are positively correlated within the networks of family and friend, pointing to a sorting criterion that is socially or trust related. These and other authors (Kremer and Miguel) also find 'negative social effects' on adoption which are attributed to strategic considerations. However, negative effects of increases in adoption within the social group could also be linked to conflicts in information rather than endogenous behavior.

3. Model Formulation

This section examines the decision of a farmer *i*, who chooses a variety *j* at time *t*, from a fixed choice set *C*. Each farmer *i* is *located* in an information region *N* which consists of trial information about a group of farmers I_N who are *close* to him. A farmer's "location" is defined by characteristics that he considers relevant for his production performance and closeness refers to the similarity along these characteristics or dimensions.

Each farmer perfectly observes the input price for each variety, P_{ijt} and imperfectly observes a set of traits, X_{ijt}. Traits in this set considered by other authors are input-use (Foster and Rosenzweig (1995) and Udry and Conley (2004)), yield (Feder and O'Mara(1981), Fisher, Arnold and Gibbs (1996)) and profitability of the new seed (Besley and Case(1994)). Indeed, farmers take decisions based on these and other imperfectly observed traits of the technology such as its complexity or its environmental sustainability. In what follows, I will refer to a generic trait or a set of traits.

In the spirit of a complicated environment and one where agents who have less than perfect cognition, the model proposes that farmers do not think in terms of full distributions; thus, they do not form expectations about the statistical dispersion of a random trait, they rather try to locate its mean value through naturally observed signal frequencies and rules of thumb (An approach to expectations that involves no probability explicitly is more simple and general than is implied by the standard treatment of expectation defined in terms of probability. Such an approach has been developed by, for example, Krantz et. al. (1971)[5]). The rules of thumb used by agents reflect that, in building their expectations about new technologies, farmers care first about the qualitative nature of observed information (particularly the closeness or relevance to their *location*), and then consider the quantitative performance.

3.1 Uncertain Traits

The uncertainty related to imperfectly observed traits originates both in their direct experience with the crop and the information they access. Imperfect observability arises for farmers who have previously used or experimented with a variety, because they might not perfectly perceive the received trait. As an illustration, farmers might not discover borers in a good share of plants and might thus perceive incorrectly the pest control performance of a new technology. Similarly, farmers might not calculate how much of a yield advantage/disadvantage a variety provides until infestation levels are really high. Also, different seeds of the same variety might perform better than others in a random manner (there is attribute variability about some mean attribute level which reflects "inherent product variability" (Roberts and Urban (1988), Erdem and Keane (1996, 1997)). The uncertainty can be described as follows. If X^N_j is the mean attribute level of variety j in location N

[5] Whether expectation or probability is the more fundamental concept has been a subject of debate in the theory of subjective probability. Krantz et. al. (1971), however, give axioms for conditional expected utility directly -- expected utility is not analyzed into utility and probability in the standard fashion where utility is the integral (sum in discrete case) over all states of the world of the products of probability and utility in each state-- and show that structural axioms are required to prove that the representation in terms of probability is possible.

$$X^N_{jt} = X^N_j + \delta_{ijt} = X_j + \xi^N + \delta_{ijt} = X_j + \gamma_{ijt} \qquad \xi^N \sim \text{d1}(0, \boldsymbol{sz^2}),\ \delta_{ijt} \sim \text{d2}(0, \boldsymbol{sd^2}) \qquad (1)$$

where δ_{ij} refers to the variability of farmers experiences with the attribute levels of varieties and ξ^N to the variability of these experiences with respect to the overall mean. The respective variances of these random terms are sd^2 and $\boldsymbol{sz^2}$.

Farmers who have not grown variety *j* in the past depend on trait information based on use or trials on others' farms (received through neighbors, media, seed dealers, university extension agents, and so forth). Thus, they receive H *noisy* signals from different sources:

$$O^h_{ijt} = X_j + \omega^h_{ijt} \qquad \omega_{ijt} \sim \text{d}(0, \boldsymbol{sw^{h2}})\ ,\ h=1,\ldots\text{H} \qquad (2)$$

where ω^h_{ijt} is the noise of information source *h* and sw^2 is the variance of ω^h_{ijt} and $sw^2 > sd^2 + sz^2$.

Farmers are imperfectly informed and hence uncertain about new varieties' mean trait levels in their location X^N_j. Furthermore, they do not observe the distribution or even variance of the different types of information noise (sz^2, sd^2, sw^{h2}). They form expectations or beliefs about variety mean trait levels (only) based on observed information up to time *t* and on subjective perceptions.

For each variety $j \in C$, let $E_N(X_j)$ be the farmer's expected mean trait level for variety *j* in location or region *N*. Let $I_i(t)$ be the information set of farmer *i* which contains all information that he has *observed* up to time *t* about the factors that affect his expected utility. The "state" of a farmer consists of all factors that affect his expected utility. If these factors are all observed, then $I_i(t)$ characterizes the state of a farmer. If there exist unobserved factors and all information about them is contained in $\Delta_i(t)$, then the belief about the mean $E_N(X_j)$ might differ from $E_N(X_j | I_i(t))$, the belief about mean traits conditional on the information that actually is observed by *i*. That is, the farmer recognizes that the information that is accessible to him is not necessarily a good predictor of performance under his own growing conditions. More specifically:

$$E_N(X_j) = E_N(X_j | I_i(t)) * \psi_j(\{|\text{I}_i(\text{t})|_{1,\ldots}, |\text{I}_i(\text{t})|_{\text{J}}\}, e_i) + E_N[X_j | \Delta_i(t)] * \psi_j(\Delta_i(t), e_i) \qquad (3)$$

where $\psi_j(\{|I_i(t)|_{j1,\ldots}, |I_i(t)|_{jJ}\}, e_i)$ is the likelihood that a farmer assigns to observable information about new varieties being an accurate predictor of his own situation [for simplicity of notation I will onwards denote $\psi_j(\{|I_i(t)|_{j1,\ldots}, |I_i(t)|_{jJ}\}, e_i) = \psi_j(.)$) and $\psi_j(\Delta_i(t), e_i)$=1- $\psi_j(.)$]. This likelihood depends on the amount of information about each specific variety in *i's* information set, $|I_i(t)|_k$, and on *i's* computational ability, e_i, in the following way: $\psi_k'(.)\geq 0$, $\psi_{kk}''(.)\leq 0$, where *k* can be e_i or $|I_i(t)|_j$. Also, if alternative varieties can be classified according to their similarity to *j*, such that *g1* are all similar varieties and *g2* are not, then $\psi_{|Ii(t)|j}'(.) > \psi_{|Ii(t)|g1}'(.) > \psi_{|Ii(t)|g2}'(.)$.

Then, the beliefs of the farmer about the mean trait levels of a variety are *confidence-weighted-averages* of one objectively rational and one purely subjective part. The objective part relates to his information set. The subjective part is based on perceptions or hypotheses about how unobservable factors might affect the local mean traits. It is viewed as purely subjective in the sense that it does not depend on observable information. Further implications of (3) will be discussed later in this section.

3.2 Utility

The local utility that a farmer *i* gets from a crop variety *j* in period *t* depends on its perfectly and imperfectly observed traits as follows:

$$U^N_{ijt} = -aP_{ijt} + b\, X^N_{ijt} + \varepsilon_{ijt} \tag{4}$$

Equation (4) is an indirect utility function with the income term suppressed.[6] P_{ijt} is the price farmer *i* pays for variety *j*, which is perfectly observed and X_{ijt} represents the generic trait which cannot be perfectly observed by the farmer at the time he makes his decision.

The expected utility of farmer *i* associated with the indirect utility function in equation (4), for variety *j* is:

$$E_N\,[U_{ijt}] \quad = \quad -aP_{ijt} + b\, E_N\,[X_{ijt}] + \varepsilon_{ijt}$$

Thus, $$E_N\,[U_{ijt}] \quad = \quad -aP_{ijt} + b\, E_N[X_j] + \varepsilon_{ijt}\,, \quad \text{if } j \text{ is a new variety} \tag{5a}$$

[6] The income term cancels out in the logit formulation of choice probabilities.

and $E_N[U_{ijt}] = -aP_{ijt} + b\, X^N_j + \varepsilon_{ijt}$, if j is a traditional variety (5b)

The expected utility of consuming variety j is a linear function of price and expected levels of the local mean traits, which are imperfectly observed for new varieties. [7] For traditional varieties, mean trait levels are observed, and (5b) is no longer of interest in the ensuing discussion. The second equality is true because δ_{ij} and ω^h_{ijt} have zero mean and the stochastic component, ε_{ijt}, stays the same as in (3) because it is stochastic from the researcher's point of view.

Substituting (3) into (5a) yields the expected utility for new varieties based on observed and unobserved factors:

$$E_N [U_{ijt}] = -aP_{ijt} + b(E_N [X_j| I_i(\mathrm{t})] * \psi_j(.)+ E_N [X_j| \Delta_{it}]* \psi_j(\Delta_{it} ,.)) + \varepsilon_{ijt} \qquad (6)$$

In order to choose among varieties, the farmer seeks to maximize his immediate expected utility by comparing the relative expected utility provided by alternative varieties. He builds several hypotheses about how unobservable factors might affect the local mean traits and aims to calculate an expected level of the mean trait across hypothetical scenarios. The farmer knows that X^N_j takes values on the real line with $x^N_{jo} = \min\{X^N_j\}$ and $x^N_{jG} = \max\{X^N_j\}$ minimum and maximum feasible values; and that $E_N [X_j| \Delta_{it}]$ can take any value $x^N_{jg} \in [x^N_{jo}, x^N_{jG}]$. In order to estimate the most likely scenario, or average over several of them in a specific period,[8] the farmer would have to associate each specific x^N_{jg} , $g=1,..G$ with a likelihood $\mathrm{p_g}^k$, of occurrence in each hypothetical scenario $\mathrm{s_k} \in \Delta_i$, where k=1,…K and $\cup_k \mathrm{s_k} = \Delta_i$ .That is, the farmer would have to calculate $E_N [X_j| \mathrm{s_k}]= \Sigma_g\, x^N_{jg}\, \mathrm{p_g}^k(\mathrm{s_k})$, for all k whereby he would have to estimate GxK likelihood values. However, several studies suggest that, particularly in complicated environments, decision-makers neither wish nor tend to infer such probabilities (Ellsberg 1961).

Instead farmers consider a full range of potential outcomes. Thus $E_N [X_j| \Delta_{it}]$ is replaced by a vector of feasible potential realizations and $E_N [U_{ijt}]$ belongs to a set of possible expected

---

[7] By focusing on uncertainty in mean attribute levels, I assume risk neutral preferences. An analysis of risk behavior is eyond the scope of this paper. It is however likely to be an important part of decision making, especially under bounded rationality. The model can be generalized to include risk by including a quadratic term of the expectation of the attribute:

$E[U_{ij} | I_i]= -aP_{ij} + b\, E[X_{ij}| I_i] – b\, r\, E[X_{ij}| I_i]^2 - b\, r\, E[\, (X_{ij} - E[X_{ij}| I_i]\,)]^2 + e_{ij}$

[8] For simplicity of notation, I omit the period subscript of scenarios in the arguments below.

utilities. More specifically, let $R\{E_N\,[U_{ijt}]\,\}$ denote the set of possible expected utilities considered by the farmer, then:

$$E_N\,[U_{ijt}] \;\in\; R\{\,E_N\,[U_{ijt}]\,\} \text{ where,}$$

$$R\{E_N\,[U_{ijt}]\}\equiv$$

$$\{\, -aP_{ijt}+b(E_N\,[X_j|\,I_i(t)]*\psi_j(.) + x^N_{j0*}\psi_j(\Delta_{it},e_i))\;,\; -aP_{ijt}+b(E_N\,[X_j|\,I_i(t)]*\psi_j(.) + x^N_{j1*}\,\psi_j(\Delta_{it},\,e_i))\,\} \tag{7}$$

Making a decision based on such a multi-valued expected utility can be described as a choice under *ambiguity*, which originates because the farmer feels that there is no way to infer meaningful probabilities for the hypothetical scenarios determined by unobservable factors. In this case, the farmer's choices are likely not the result of a standard expected utility maximization process. Yet, if their behavior is "deliberate and orderly…it can be described in terms of a simple, specified decision rule" (Ellsberg; p.646).[9]

From the specification in Equation (7) one can deduce that the size of the range of expected utilities considered, $|R\{E_N\,[U_{ijt}]\}|$, depends on three factors:

(1) The spread of the local distribution of trait values for variety *j*,

$$Spread_j^N = (\max\{X^N_j\} - \min\{X^N_j\}) = (x^N_{j1} - x^N_{jo})$$

(2) the amount of trait information that the decision-maker has observed for each variety, $\{|I_i(t)|_{1,\ldots},\ |I_i(t)|_J\}$

(3) and his computational ability, e_i .

<u>Result 1.a</u>: For similar localities, farmers that have more experience or more data and more computational ability are also more confident about this observed data and have less uncertainty about the utility received from new technologies.

<u>Result 1.b</u>: The degree of (unmeasurable) uncertainty that farmers with similar information sets and computational ability face increases in the spread of the local trait distribution.

[9] The same author points out that "In reaching a decision under these circumstances, many people seem to act conservatively. Without actually expecting the worst, they choose to act as if the worst outcomes were somewhat more likely than their best estimates of likelihood would indicate".

<u>*Result 2:*</u> An individual's *comparative ignorance*[10] might influence choices when faced with a complicated environment. That is, an *ambiguity averse* farmer --who maximizes his expected utility in the worst case scenario)-- and who chooses among two new varieties that cost the same and have the same based-on-observable-information expected trait means, will choose the one that is more familiar to him.

Proof of Result 2: In order to see this, one can think of the interval (min$\{X^N_j\}$, max$\{X^N_j\}$) as a normalized interval (0,1). By substituting this into the expected utility region (7) it can be seen that the uncertainty region depends only on $\psi_j(\Delta_i(t), e_i)$=1- ψ_j(.). Recalling from (3) that $\psi_{|Ii(t)|j}$'(.)> $\psi_{|Ii(t)|g1}$'(.)> $\psi_{|Ii(t)|g2}$'(.), it is obvious that this region is smaller for more familiar varieties because the farmer has more confidence about correctly observing their traits. Since P_{ijt} and E_N [X_j| I_i(t)] are the same for the compared varieties the lower bound of the expected utility region is higher for the familiar variety, and the uncertainty averse farmer makes his choice maximizing lower bounds.

Result 2 highlights the importance of the similarity notion, not only when sorting out information signals for calculating expected mean traits, but also for reducing the range of uncertainty that relates to ambiguity. It reflects the fact that choices by farmers might be based on familiarity and/or popularity (because the size of the information set that a farmer has for a specific variety *j* depends on the number of observations that a farmer receives from own experimentation and from other's experimentation) not only of the chosen variety but of similar varieties. This explains the type of choice "inertia" (E&F) that has been observed by several studies of agricultural technology adoption, but also points to the possibility of "transfer" (Suppes 1994) across similar varieties. It thus points to a fundamental weakness of previous studies, which have focused on the adoption of a single new variety contrasting it to a fundamentally different traditional alternative, which is not able to capture this transfer. As an example, a farmer who already grows non-genetically-modified herbicide-resistant (HR) varieties might be more likely to transfer to new, never used, genetically-modified (GM) HR varieties than a farmer who never chose HR existing varieties and who also has no experience with the new GM types. The former faces lower uncertainty regarding the management of the new varieties. A study that does not consider this type of similarity and which uses own

[10] See Fox and Tversky in Shafir (2004).

experience as a determinant of demand for the new variety will assign the same likelihood of adopting the new GM-HR variety to both farmers.

3.3 Identification Issues

A researcher who seeks to estimate expected utility in equation (6) and who relies on observational data faces two *identification* issues, that combined create a third identification issue:

Identification Issue 1: Unobservability of the second term in (6), $E_N\ [X_j|\ \Delta_{it}]$

Identification Issue 2: Unobservability of the farmers information region, or reference information set N and of its boundries.

Identification Issue 3: Because of identification issues 1 and 2, the trait preference parameters, b, might be not identified or be only partially identified.

In this section I only elaborate on the first and third identification problems and derive some measurement biases that arise in standard frameworks because of them. I discuss the second identification issue later in the farmer-location section.

The first identification problem is related to the fact that, given a random sample and observability by the analyst of I_i(t), the sampling process asymptotically reveals $E[X_j|\ I_i(t)]$, but is uninformative about $E[X_j|\ \Delta_i]$. Thus, based on the empirical evidence alone, farmers' expectations about the mean trait level $E[X_j]$ cannot be *identified* by the researcher, even conditional on all observed information.

By the nature of $E[X_j|\ \Delta_i]$ no observable variables are useful instruments for precisely predicting this expectation. Thus, researchers seeking to identify $E[U_{ijt}]$ have assumed that although agents are unable to observe some factors influencing their decision, they do observe their distribution. Thus, they observe $E[U_{ijt}]= bE[X_j|\ I_i(t)]+v_i$. The analyst then specifies, somewhat arbitrarily, a distribution for v_i across the population, which allows both the decision maker and the analyst to infer Ev_i and the beliefs $E[X_j]$ from the actual choices. However, this identification approach is problematic, not only because it presupposes that individuals are able to measure the uncertainty, but also because it assumes that the degree of this uncertainty is exogenous to characteristics of the individual such as cognitive ability or experimentation.

Since $E[X_j|\ \Delta_i]$ can lay anywhere between $[x^N_{jo}, x^N_{jG}]$, the decision-maker uses some heuristic method that seems sensible for the setting under consideration in order to decide among alternatives. The most common approach of researchers has been to assume that the decision-maker is *ambiguity averse* and accepts the greatest expected value of his uncertainty horizon which will give him some minimal specified utility level. Since this minimal level is not known by the researcher, the practice is to derive a set of possible results based on several thresholds (How specifically this approach determines choice is examined in the Decision Rules section below). Since expectations determine choices and choices reveal preferences for traits, identification problems related to expectations lead to identification problems for preference parameters.[11]

Result 3 (Identification Issue 3): Expectation-identification issues lead to the following identification problems for preferences for traits:

1.) For given choices, a range of possible expected utilities associated with the unobservability of $E[X_j| \Delta_i]$ determines a range of possible preferences for traits.
2.) The variation in expectations associated with location characteristics might be confused with preference heterogeneity. This might result in incorrect willingness-to-pay estimates for choice traits.

Proof of Result 3.2.: Let Z_i be the characteristic that determines a farmer's location, such that $E_N [X_j] = Z_i E[X_j]$. Then the corresponding expected utility is $E_N [Uijt] = -\alpha p + bE_N [X_j] = -\alpha p + bZ_i E[X_j]$. If the expected willingness-to-pay for a trait is defined by EWTP(x) = dp/d(E_N x), then EWTP(x) = $- bZ_i/\alpha$.

Analysis that tests for heterogeneity in preference parameters across individuals commonly specifies preference heterogeneity as $b\hat{}=bZ_i$. If the researcher does not control for location he will find an heterogeneous measure of willingness-to-pay for traits that equates in form the one above (EWTP(x) = $- bZ_i/\alpha$) but he will justify it on completely different grounds. This is so because the expected utility will be identical

$E [Uijt] = -\alpha p + bZ_i E[X_j]$ in the two settings.

Alternatively, if expected willingness to pay is defined locally as E_N WTP(x) = dp/d(E_N x), then an analysis of preference heterogeneity that assumes $E[X_j] = E_N[X_j]$ will incorrectly estimate E_N WTP(x) as $- bZ_i/\alpha$.

3.4 Measurement Biases Under Full Rationality

[11] This point was noted as early as in 1961 by Ellsberg.

The implicit identification approach, the one most commonly used, views farmers as having a 'very big' information set and perfect cognition or computational ability. If this is the case, and the farmer has *full observability*, his information set would contain all information that might possibly influence his expected utility and/or the probability distribution of the future expected utilities (such that this set describes his state), and then $\psi_j(\{|I_i(t)|_{1,\ldots}, |I_i(t)|_J\}, e_i) = 1$.

Furthermore, the farmer is generally assumed to compute his expectations based on all information rather than on those expectations that arise from conditions most similar to his. The assumption underlying this reasoning is that the process determining trait outcomes is stationary (over time and across space), which implies that the true mean trait should not change across locations. This, in the world of subjective probabilities, is further tied to the assumption of exchangeability in the data. Data or information about a stationary process is most likely to be exchangeable (See Suppes (1994) for a discussion on these concepts). That is, the order (in terms of location or timing) of the data is not relevant for determining its expectation, or as Epstein (2006) puts it, information about each observation is assumed to be *a priori* the same.[12]

In terms of the model presented, these assumptions result in a simplified form of (6):

$$E_N [U_{ijt}] = E [U_{ijt} \mid I_i(\mathrm{t})] = -aP_{ijt} + b\, E[X_j \mid I_i(\mathrm{t})] + \varepsilon_{ijt} \qquad (6')$$

This approach has been a standard assumption in the learning literature because it is a way to get back to the world that the researcher is able to observe. However, estimations based on these assumptions are subject to two sources of measurement bias (MB) in a bounded-rationality environment:

$$MB1 = (E_N [X_j \mid I_i(\mathrm{t})] - E[X_j \mid I_i(\mathrm{t})])$$

$$MB2 = (E_N [X_j] - E_N [X_j \mid I_i(\mathrm{t})])$$

The first source of measurement bias arises because the farmer does not estimate a mean trait based on all observed signals. The farmer cares about the quality of these data (in terms of the

[12] Exchangeability is a weaker form of independence that is used by subjectivists and is assumed by the Bayesian framework (Suppes 1986).

information that they convey regarding potential outcomes on his own farm). Therefore, he sorts out the information he receives before processing it. As opposed to standard models of agricultural technology adoption, this model emphasizes the need to sort out observed information, not only the need to access information, as an obstacle to learning about true traits of a new technology.

Result 4 (Measurement Bias under full rationality 1): The assumption of data exchangeability in standard frameworks, results in lack of attention to the information sorting and processing undertaken by farmers.

Proof of Result 4 : The data in this model are composed by trial observations. Exchangeability of observations across space implies that location is not relevant to the trial outcomes. Thus $E_N [X_j \mid I_i(\mathrm{t})] = E[X_j \mid I_i(\mathrm{t})]$, implying that the data do not need to be sorted out. The next section provides a deeper analysis of the sorting process.

The second source of measurement bias is related to the perception of the farmer that the data he has observed in his location are non-random (for a given location). That is, that information signals are not based on trials that are representative of the whole local population such that when randomly drawing from them and calculating a mean trait, this mean will not necessarily be close to the true local mean.

Result 5 (Measurement Bias under full rationality 2): The measurement bias associated with the assumption of full observability of the distribution of outcomes in a specific location is larger for: a) farmers with lower computational ability, b) farmers with lower amount of experience with the variety or with closely related varieties, c) farmers whose farming conditions differ mostly from the conditions of previous variety *j* adopters.

Proof of Result 5 : The measurement bias 1 is defined as MB1 $= E_N [X_{ijt}] - E_N [X_{ijt} \mid I_i(\mathrm{t})]$. Substituting equation (3) yields:

$$MB1 = E_N [X_j| I_i(\mathrm{t})] * \psi_j(.) + E_N [X_j| \Delta_{it}] * \psi_j(\Delta_{it} ,.) - E_N [X_j| I_i(\mathrm{t})]$$

$$= E_N [X_j| I_i(\mathrm{t})] * (\psi_j(.) - 1) + E_N [X_j| \Delta_{it}] * \psi_j(\Delta_{it} ,.)$$

$$= E_N [X_j| I_i(\mathrm{t})] * \psi_j(.) - E_N [X_j| \Delta_{it}] * \psi_j(.) + E_N [X_j| \Delta_{it}] - E_N [X_j| I_i(\mathrm{t})]$$

$$= (E_N [X_j| I_i(\mathrm{t})] - E_N [X_j| \Delta_{it}]) * \psi_j(.) - (E_N [X_j| I_i(\mathrm{t})] - E_N [X_j| \Delta_{it}])$$

$$= (E_N [X_j| I_i(\mathrm{t})] - E_N [X_j| \Delta_{it}]) * (\psi_j(.) - 1)$$

$$= (E_N [X_j| \Delta_{it}] - E_N [X_j| I_i(\mathrm{t})]) * (1-\psi_j(\{|\mathrm{I}_i(\mathrm{t})|_{1,\ldots}, |\mathrm{I}_i(\mathrm{t})|_{\mathrm{J}}\}, e_i))$$

In order to better understand the second source of measurement bias it is necessary to uncover the complexity of the information sorting and processing undertaken by decision-makers. This data selection process amounts to contradicting the exchangeability assumption of standard frameworks and reveals the difference between $E_N[X_j| I_i(\text{t})]$ and $E[X_j| I_i(\text{t})]$ for the farmer. This will be discussed in the next section, where the expectation formation process $E_N[X_j| I_i(\text{t})]$ is described.

3.5 Expectation Formation Process

A farmer's current beliefs about the observable local mean traits of a variety *j* can be partitioned according to the time of information acquisition. At the beginning of a period, before making choices among varieties, he has some prior expectation that is based on observed signals up to that time, $E_N [X_j |I_i(t-1)]$. At the end of the period, the harvest is done and the farmer observes the outcomes on his land and hears about the outcomes of other farmers. He observes the difference between actual trial realizations X^h_{ijt} and his *ex-ante* calculations and uses this difference to infer part of the mean trait (the observable part) in the next period ($E_N [X_j | I_i(t)]$), as follows.[13] For simplicity of notation, I will include own experience as part of the set *H* of information sources and denote the information received from each source by X^h.

$$E_N [X_j | I_i(t)] = E_N [X_j |I_i(t-1)] + \Sigma_h \mathrm{D}^{h}_{ijt}\beta^{h}_{ijt}*(X^{Nh}_{ijt} - E_N [X^{N,h}_{ijt} |I_i(t-1)]) \qquad (8)$$

where D is an indicator that is equal to one if farmer *i* has received information from source *h* about variety *j* and at time *t* , β^h are weights attached to the surprise elements of different information sources ($0\leq \beta^h\leq 1$). Since the random terms ω^h_{ijt} (in equations 1 and 2) have zero mean the expected level of information obtained via experience is equal to the expected level of information received via other information sources: $E_N [\mathrm{X}^{N,h}_{ijt} | I_i(t-1)] = E_N [\mathrm{X}^{N}_{j} | I_i(t-1)]$ and (8) can be rewritten as:

$$E_N[X_j | I_i(t)] = \rho_{o\,ijt} *E_N [X_j |I_i(t-1)] + \Sigma_h \rho^h_{1,ijt}*X^{N,h}_{ijt} \qquad (8')$$

where $h= 1,...,H$, $\rho_{o\,ijt} = (1- \Sigma_h\rho^h_{1,ijt})$, $\rho^h_{1,ijt}= \mathrm{D}^{h}_{ijt}\beta^{h}_{ijt}$

[13] However, this is reevaluated right before making the next choice as explained below.

The ρ_i coefficients can be interpreted as the weights farmers attach to alternative information sources in updating their expectations about variety traits. Recursively solving equation (8'), it can be rewritten as an expression that depends on the accumulation of information in every period:

$$E_N[X_j \mid I_i(t)] = \rho^{\wedge}_{o\,ij} * E_N[X_j \mid I_i(0)] + \Sigma_h \rho^{\wedge h}_{1,ij1} * X^{N,h}_{ij1} + \ldots + \Sigma_h \rho^{\wedge h}_{1,ijt} * X^{N,h}_{ijt} \qquad (8'')$$

where $\rho^{\wedge}_{o\,ij} = \prod^{t}_{k=0} \rho_{o\,ijk}$, $\rho^{\wedge h}_{1,ij1} = \prod^{t-1}_{k=0} \rho^{k+1}_{o\,ijk}(1- \rho_{o\,ij1})$, …, $\rho^{\wedge h}_{1,ijt} = (1- \rho_{o\,ijt})$,

Farmers can obtain maximum one observation from own experience each year. However, other sources of information provide multiple signals (O^{h}_{ijt}) that are sorted out and processed by the farmer in terms of their relevance to his own potential outcomes. [14] In order to do this, the farmer first considers his *location* in terms of the observable factors that are most relevant for his trait outcomes. Let farmer *i*'s perceived location be denoted by the vector of farming conditions Z_i =(*Gi, Li, Mi*), where *Gi* is farmer *i*'s geographical location, *Li* are farm characteristics, and *Mi* is a farmer's management ability. The farmer is partially informed about the location or trial conditions on other farms Z^h_l , $l=\{1,\ldots,n\}: l \neq I$, on which each information signal O^{h}_{ijt} has been based. He evaluates the distance between his own relevant farming conditions Z_i and the observed Z^h_l 's, using a distance function $d(Z_i\, Z^h_l)$. He then selects a set of signals $N = \{O^{h}_{ijt} : d(Z_i\, Z^h_l) \geq d_{ki} \}$ that are close to him (In other words, his similarity function is given by $s(Z_i\, Z^h_l)=I(d(Z_i\, Z^h_l) \geq d_{ki}))$. Thus, if d is above a certain threshold for a specific O^{h}_{ijt} he includes it in his summary measure of others' information.

How might he construct such a summary measure? Two modes of reasoning have been studied in the financial investment literature and are very common in generating these types of assessments (Gayer, Gilboa and Liberman 2004). According to these, the farmer might use a case-based or a rule-based summary measure. The former relies on an (weighted) average of signals that considers locations through the similarity or distance measure. The latter uses each of the factors determining the location to infer a rule about the relationship of factors to outcomes, and then uses this rule to predict outcomes in the specific location.

The case-based assessment is:

[14] Recall that in equation (8'') I used simplified notation such that $X^{N,h}_{ijt}$ includes own and others' signals.

$O^{N}_{jt} = \Sigma_l \Sigma_h \, s(Z_i \, Z^h_l) O^{h}_{ijt} * (\Sigma_l \Sigma_h \, s(Z_i \, Z^h_l))^{-1}$ and given the similarity function considered

above

$$O^{N}_{jt} = \Sigma_l \Sigma_h \, O^{h}_{ijt} * (N+H)^{-1} \tag{9}$$

The rule-based assessment is:

$$O_{ljt} = \alpha Z^h_l = \alpha_1 \, G_l + \alpha_2 \, L_l + \alpha_3 \, M_l + e_l$$

such that, $O^{N}_{jt} = \alpha_1 \, Gi + \alpha_2 \, Li + \alpha_3 \, Mi$ (10)

A farmer using method one averages all the “cases” that have similar characteristics tohis own. A farmer using the second method is more sophisticated, weights each of the relevant characteristics in terms of their influence on outcomes, and uses all stored information in order to approximate outcomes in his location. While both farmers seek to approximate outcomes at their own location, the former believes that outcomes at faraway locations are irrelevant to his decisions. Which type of method the farmer uses is an empirical question.

3.6 Farmer Location Discussion

The second aspect mentioned above which generally poses a problem to the researcher is the identification of a farmer’s “information region” *N*. Information flows have been regarded as the result of communication with others through e.g. family-, ethnic-, religious and/or financial networks. As an example, Munshi and Myaux (1998) explain how in rural India communication barriers may arise across religious groups. Thus, the identification problem has been generally seen as the problem of distinguishing between the effects on behavior of flows of information between farmers along those dimensions and the effects of spatially-correlated growing conditions. However, in a world with very widespread access to information (where new information technologies such as internet, television and cellular phones have reached remote rural areas), the information available to farmers across the dimensions mentioned above might look very similar. Arguably, what matters in such a context is which information the farmer stores and how he processes it. The perception of a farmer about his *location* determines the dimensions that a farmer considers when sorting out available information; that is, the observable factors that he thinks are most important for the outcomes in his farm. The distance or *similarity* measure used to define his neighborhood is then a function of these factors. In agricultural settings, location is based on farming conditions which might be determined by variables like geographical location, farm characteristics, and/or farmer management ability.

Simple approaches have assumed that the farmer's perception of location is associated with his farm's geographical location only. In agricultural settings this might be sensible given that crop performance is tightly related to agronomic conditions and given that information disseminated about trials often includes the geographic location but no specifics about plot soil, size or management conditions. In this case, flows of information and local growing conditions are perfectly correlated and their effects cannot be identified separately. The identification problem for the researcher is identifying the reach or the limits of the geographical location (geographical window width). However, if farmers are able to control for other dimensions when building the information region these should be considered as well and this is done as in equations (9) and (10). In this case, the identification problem for the researcher is identifying which characteristics are observed in order to characterize his location.

In period zero, when a new variety is introduced and nobody has experimented with a new variety, a farmer cannot approximate local performance based on the above rules. Thus, his prior is that the mean trait level of that variety is normally distributed about the crop class local mean (A_N). Thus, letting $I_i(0)$ denote the farmer's prior information about the crop class, we have that $E_N\,[Xj\;|Ii(0)] = A_N$. Notice that a farmer's initial belief is associated with his location, which is determined by his farming conditions. Thus, initial beliefs are not only crop specific but take into account farming conditions. Sociologists usually think of these initial reference values as reflecting different "adoption thresholds" or as differences in responsiveness to social pressure (Valente 1995; 2005; Young 2006).

Substituting initial beliefs and farmers location into (8'') yields:

$$E_N[X_j \mid I_i(t)] = \hat{\rho}_{o\,ij} * A_N + \ldots + \hat{\rho}^{own}_{1,ijt} * X^N_{ijt} + \hat{\rho}^{others}_{1,ijt} * O^N_{jt} \qquad (8''')$$

Equations 1-8 define a general model which by itself does not preclude a Bayesian interpretation and implementation of the same. The key aspects which further determine the informational structure of the individuals, apart from the nature of the information region (local or not), is the definition of the weights attached to different information types (new versus old and own versus others') and the timing of events in each period.

3.7 Making the Model Operational

The Bayesian updating mechanism has been the most widely used expectation updating framework in the technology adoption studies that implements dynamic adoption frameworks (Feder and O'Mara (1981), Besley and Case (1994), Foster and Rosenzweig (1995)).[15] It provides important intuition about underlying relations of factors that affect learning under a specific setting; that is, when the assumptions of the model hold. However, they are strong assumptions, both in terms of the decision maker's cognitive ability and the resulting aggregate prediction pattern (Young 2006, Fisher, Arnold and Gibbs 1996). They might have very unrealistic implications.

In this section I first explain the main assumptions of the Bayesian framework and then justify and propose an alternative approach. A discussion follows explaining how looking at the problem through the Bayesian lens brings the researcher back to a reliance on the full rationality assumption. I argue that the restrictiveness of this approach fundamentally arises from the assumptions required to pin down the weights that decision-makers attach to different types of information. Another fundamental weakness of the approach, which is not related to its restrictiveness, is the lack of guidance as to how prior beliefs are generated. This weakness has lead to a misconception about the importance of the initial beliefs of agents in determining adoption and has lead to a *selection bias* problem in past studies of adoption.

<u>Bayesian Updating</u>

A standard Bayesian framework assumes that the error terms in equations (1) and (2) are *iid*, normally distributed and that the trait process is stationary (the true mean, variance and autocorrelation structure do not change over time). Then, a *Bayesian* updating rule can be used. This rule says that the information weights β^h in equation 8 are functions of the noise of new information (variance s^{h2}, including sd^2 of own experience and sw^{h2} of alternative sources) adjusted by the precision of the beliefs (the inverse of the variance of the individual's prediction error, $1/s^{x2}(t)$). All of the former parameters are assumed to be known by the farmer who assigns

[15] Exceptions to this are Ellison and Fudenberg (1993) and Manski (1993). The former considers boundedly rational agents and an updating mechanism based on popularity weights and inertia. The latter sees decision-makers applying non-parametric regression methods.

the following weights to alternative information sources (the other terms are as explained in the previous section):

$$\beta^{h}_{ij}(t) = [1 + s^{h2}/s^{x2}_{ij}(t)]^{-1} \qquad (10)$$

where

$$s^{x2}_{ij}(t) = [1/ s^{x2}(0) + \Sigma_h(\Sigma_t D^{h}_{ijt}/ s^{h2})]^{-1} \qquad (11)$$

is the variance of the prediction error, [16] which is a function of *exchangeable* information accumulated over time through different sources, the initial prediction error, and the precision of information sources.[17] Notice --in equations 8-8'''-- that the weights assigned to the different information sources vary over time because the variance of the prediction error s^{x2} varies over time. More specifically, each piece of observed information reduces the prediction error by the same amount, and regardless of the moment in time in which it was observed (independence of observations is then an obvious requirement. If observations were strongly dependent, then many observations would count no more than a single one).

Not only must the data received by individuals be *iid* normally distributed, but individuals are assumed to know this and be able to characterize the true distribution of the information error as well. Additionally, these individuals keep track of all the trait levels observed historically, not only on their farm but also on their neighbors' farms and remember for their calculations, at the time of their decision, all outcomes and advice from neighbors, seed dealers, extension agents and advertising seen in the media, since the new technology was introduced. Then, their information set is big enough to infer the past distribution of outcomes. Additionally, in order to make inference about the prediction variance in every period, the farmer is assumed to combine information signals through a process that amounts to a linear regression

---

[16] Since each signal is considered unbiased by the farmer, the best minimum variance way of combine the information is: $X^{\wedge}= X_j*\mathbf{1}+ (X^{\wedge}- EX) = X_j*\mathbf{1}+ u$ where X^ is a vector of all realizations observed by the farmer and u a vector of prediction errors of the same dimension. Thus, V(u) is a (H+1)*(H+1) diagonal matrix with terms sx2(0), $s^{1,2}/n_{1t},\ldots, s^{H,2}/n_{Ht}$ terms in the diagonal. Using the Gauss-Markov Theorem $s^{x2}_{ij}(t) =(\mathbf{1'} V(u)^{-1}\mathbf{1})^{-1}$ which, for uncorrelated errors, is equation (7).

[17] In the seminal study by Foster and Rosenzweig (1995) utility is assumed to depend directly on the variance of the input-use prediction error. That is, the relevant trait is not input-use but its variability.

(minimizing the squared error of the linear system of information signals) and invert an (H+1) squared variance-covariance matrix to finally know the variance of the prediction.

According to (8, 10, 11), $E_N\,[X_j \,|\, I_i(t)]$ is a function of the farmer's full prior distribution over Xj at time t. An important underlying assumption which is necessary to obtain this tractable and interpretable form of belief is that the process of the realized traits is stationary and it is also to be perceived so by the decision-maker. Additionally, it has been the norm to assume that common observations are available to decision-makers (and to the researcher) and these commonly accepted observations will often force their beliefs to converge (Suppes 1986).

4. Alternative Approach

When the complexity of the environment is perceived by the decision-makers, they might perceive information signals (equations 1 and 2) as not being independent of each other, or identically distributed (Epstein and Schneider (2006) describe such a setting, where agents doubt that their observed data are identically distributed). This alone renders the standard Bayesian updating inapplicable (and inappropriate) for this situation Although some alternatives to the standard Bayesian framework have been recently studied by Gilboa and Schmeidler (1989), Schmeidler (1989) and Epstein and Schneider (2003; 2006), they generalize the framework in a way that might require even more cognitive ability of decision makers and contrasts with increasing evidence found by psychologists regarding subjective assessment of probabilities and expectations (Kahneman and Tversky (1973; 74), Slovic et. al. 1982; Shafir 2003). [18]

Those generalized frameworks, as well as the standard Bayesian, take probabilities as primitives. Thus, even if a decision maker is allowed to be unsure and face ambiguity, he is assumed to think of the world like a statistician, in terms of probabilistic representations, with underlying coherent probability measures in probability spaces with a specific ordering and algebra of events. Not only is it hard to think in those terms, but even if one is a statistician it is

[18] These authors have axiomatized behavior of agents in environments that are not well understood by agents. They have modeled agents decision-making and learning under ambiguity in two different ways. One has been to assume that agents have multiple priors and that they use likelihood-ratio test procedures to decide among priors. These agents are assumed to, then, only update the chosen priors (the ones that fit the observed data best) possibly in a Bayesian manner. A type of non-Bayesian updating has been studied by Epstein (xxx), but it requires high sophistication of the agents in order to be able to learn. The second one consists of specifying a non-additive measure of probability for the different types of uncertainty and using a non-standard integration procedure (based on Choquet integrals) to estimate expected utility.

hard to prove the existence of this underlying probabilistic representation or to test what axioms in fact are satisfied. A statement by Patrick Suppes provides an argument against an interpretation of subjective expectations in terms of subjective probabilities:

> "It may be argued that in ordinary experience expectation rather than probability is the more widely used concept. As I have put it elsewhere, the argument for this is evident from a practical standpoint. Once we leave events and talk about what correspond to random variables it is natural in ordinary talk to want to know only the expectation and not the full probability distribution. Thus we talk about the expectation of walking at least 10 km in the next 3 days, the expectation of at least 3 cm of rain in the next 8 hours, or the expectation that the rate of inflation in the next year will be about 9%. In all of these cases, we are dealing in a natural way with a quantitative variable, but we are not prepared to give, and are really not interested in giving, the full probability distribution of that variable or even, usually, the variance." (Comment on *The Axioms of Subjective Probability*, P. Suppes, P.59)

I, thus, choose to use the simple but general updating framework presented above (without Bayesian updating). It is general in the sense that the farmer is neither assumed to characterize the whole probability distribution of the random variable X^N_{ijt} , nor to think in terms of probabilities. Thus, a belief is conceived as a mean belief or an empirical frequency, rather than as a full description of likelihood under all alternative states of the world. As such, no sample space or algebra of events is involved.

Up to now, observability and computational ability issues have been partially addressed by the proposed use of local expectations. This is similar in nature to the idea of decision-makers being able to observe outcomes within their "window width"; as proposed by E&F. It limits the information considered by decision-makers and addresses informational constraints, computational capacity problems or simply farmer's beliefs regarding the irrelevance, to their own decisions, of experiences at faraway locations. Also, in equation (1) own-information signals were allowed to have a local error component that is distributed differently from the idiosyncratic component.

There are three key aspects further representing alternatives to the standard framework which remain unspecified (or just briefly mentioned). One is the definition of the weights attached to different information types in each period ('new versus old and own versus others'), the second is the timing of the complete expectation formation process, and the third is the decision rule used. In the following subsections I address these issues.

4.1 Information Weights and Rules of Thumb

Farmers who do not store and process all the information that they observe care about information precision in terms of its *representativeness* --similarity to their own conditions—(Tversky and Kahneman 1974) rather than about an absolute or objective measure of variance of the information error over the whole population ("Objective" here means analysis of a signal out of context and without regard to the recipient). In fact, empirical studies have consistently found that, when available, recent own-farm information is the most precise and unbiased information available to an individual farmer" (Marra et. al 2001; Foster and Rosenzweig). Thus, when this type of information is available in a specific period, a farmer sticks to a rule that overwrites his initial beliefs and the information received through alternative sources with own experience. Since this information is scarce and accurate, over time he stores all of these signals. However, if there is one period in which the farmer decides not to grow the same variety, then in the next period he will use the most recent information received from other sources (if available) combined with his own experience in preceding periods (he does not use the initial beliefs anymore). If a farmer has never experimented with a new variety, but he has received information signals from other sources about its traits, he might use the most relevant information recently obtained from alternative sources or hold on to his prior beliefs. Kahneman and Tversky (1973) found that subjects disregard prior probabilities when they have other information, even if this information is uninformative or leads to prediction errors. They correctly utilized prior probabilities only when no other information was available. However, Wells and Harvey (1977) suggest that new and old information might conflict with each other and either source of information might be dismissed. When initial beliefs are given priority, farmers have been said to be *conservative.* Which source of information is used in this case could be tested empirically in terms of the predictive ability of the model. Since information signals from others are less accurate and more abundant, a farmer only stores the most recent ones received. Finally, if the farmer has not experimented recently with a specific variety and neither have any of the farmers close to him, the farmer has no other source of information than his initial belief.

This rule of thumb description implies that over time agents separate in heterogeneous information structure categories. These categories can be summarized in terms of the different weighting schemes in Equation (8'''):

Category 1: $\Sigma_t\ \hat{\rho}^{own}_{1,ijt} = 1$ if in the last period, $D^{own}_{ijt} = 1$

Category 2: $\hat{\rho}^{others}_{1,ijt} + \Sigma_{k=1}^{t-1}\ \hat{\rho}^{own}_{1,ijk} = 1$ if for some period(s) up to t-1, $D^{own}_{ijt} = 1$

but in the last period $D^{own}_{ijt} \neq 1$ and $D^{others}_{ijt} = 1$

Category 3: $\Sigma_{k=1}^{t-1}\ \hat{\rho}^{own}_{1,ijk} = 1$ if for some period(s) up to t-1, $D^{own}_{ijt} = 1$

and in the last period $D^{own}_{ijt} \neq 1$ and $D^{others}_{ijt} \neq 1$

Category 4: $\hat{\rho}^{others}_{1,ijt} = 1$ or $\hat{\rho}_{o\,ij} = 1$ if, for <u>no</u> period up to t-1, $D^{own}_{ijt} = 1$

and in the last period $D^{others}_{ijt} = 1$

Category 5: $\hat{\rho}_{o\,ij} = 1$ if, for <u>no</u> period up to t-1, $D^{own}_{ijt} = 1$

and in the last period $D^{others}_{ijt} \neq 1$

<u>Result 6:</u> Farmers who have grown a new variety every year since its introduction rely on their own information only when forming expectations about traits. Farmers who rotate combine their previous own experience with the most recent others' experiences (as available). Farmers who have never grown a new variety might use neighbor's signals, if these are available and do not conflict with their initial beliefs. Otherwise, farmers rely exclusively on their initial expectations.

Thus, decisions among varieties are based on one of the following expected utilities corresponding to different information structures.

For categories 1 and 3:

$$E_N\,[U_{ijt}] = -aP_{ijt} + b(\psi_j(.) * \Sigma_t \hat{\rho}^{own}_{1,ijt}\, X^{N,h}_{ijt} + E_N\,[X_j |\, \Delta_{it}] * \psi_j(\Delta_{it}\,,.)) \quad (12)$$

For category 2:

$$E_N\,[U_{ijt}] = -aP_{ijt} + b(\psi_j(.) * (\hat{\rho}^{others}_{1,ijt} O^N_{jt} + \Sigma_t \hat{\rho}^{own}_{1,ijt} X^{N,h}_{ijt}) + E_N\,[X_j |\, \Delta_{it}] * \psi_j(\Delta_{it}\,,.)) \quad (13)$$

For category 4a:

$$E_N\,[U_{ijt}] = -aP_{ijt} + b(\pi_j(.) * O^N_{jt} + E_N\,[X_j |\, \Delta_{it}] * \pi_j(\Delta_{it}\,,.)) \quad (14)$$

For categories 4b and 5:

$$E_N [U_{ijt}] = -aP_{ijt} + b(\pi_j(.)*A_{N}*+ E_N [X_j| \Delta_{it}]* \pi_j(\Delta_{it} ,.)) \quad (15)$$

Defining how far in the past are the weights of the own experience signals given a non-zero value and what the relative magnitude of these weights is, precisely depends on whether the process is perceived to be stationary or not. Several processes in agriculture which determine traits like yield advantage or insecticide-savings exhibit seasonality. As an example, average levels of European Corn Borer infestations for the state of Wisconsin are shown to oscillate in intervals that peak every three to seven years since 1942 (Lambrecht 2004). Farmers who perceive this type of variation are much better off approximating the future level of infestation based on the last observed signal than based on an average over all years. Thus, in order to determine these weights a researcher should look at the type of process in question. In the same vein, recent information from very close neighbors might be a better predictor than old own information.

4.2 Timing of Events in a Specific Period

A farmer's current beliefs about the observable local mean traits of a variety *j* were partitioned according to the time of information acquisition. At the beginning of a period, before making choices among varieties, a farmer has some prior expectation that is based on observed signals up to that time. After the harvest, at the end of the period the farmer observes new information. This would correspond to a standard rational Bayesian framework if the farmer did not distrust observed information. Before making his variety choice the farmer might reevaluate the prior expectation, based on his perception about how unobservable factors might affect outcomes at his location. This corresponds to the second term in the expected trait equation presented earlier $E_N (X_j) = E_N (X_j| I_i(t))*\psi_j(\{|I_i(t)|_{1,\ldots}, |I_i(t)|_J\}, e_i) + E_N [X_j| \Delta_i(t)]* \psi_j(\Delta_i(t), e_i)$. Then, the farmer makes his choice and purchases the seed for the chosen variety.

The timing of the process described above allows for the possibility that farmers do not act according to the information provided by the observed sample information.

Result 7: Choices made based on the maximization of equation (6) do not require the farmer to be objectively or subjectively rational, in the sense that a farmer might change the ordering in which he had ranked alternative varieties earlier in the same period in a way that contradicts the preference ordering revealed by observed data.

Proof of Result 7: Let j be a new variety for which $E_N\,[X_j|\,I_i(\mathrm{t})] * \psi_j(.) - X^N{}_k > 0$, where k is a traditional variety with perfectly observed traits. Let $P_{ijt} = P_{ikt}$ and

$E_N\,[X_j|\,\Delta_{it}] > (E_N\,[X_j|\,I_i(\mathrm{t})] * \psi_j(.) - X^N{}_k)\,/(1-\psi_j(.))$, then the following two statements hold :

a) $E_N[U_{ijt}\mid I_i(\mathrm{t})] > E_N[U_{ikt}\mid I_i(\mathrm{t})]$

b) $E_N\,[U_{ijt}] < E_N\,[U_{ikt}]$

An example of the previous result in the setting of adoption of genetically modified crops is the following: a farmer has planted an insect-resistant variety in period t-1. After the harvest he receives only good signals about the mean traits of this variety, relative to regular varieties. A standard rational Bayesian farmer should act accordingly to what he has observed and continue growing the insect-resistant variety. In the present framework, the farmer might have some sudden fear regarding potential environmental damage caused by all his plantings and assign a very low expected trait value to the GM variety under this hypothetical scenario. Thus, although he weights the good news observed scenario higher, the expectation on his nightmare scenario might be so highly negative that it might cause him to disadopt and only buy a regular variety in the next period.

4.3 *Alternative Decision Rules and Estimation Strategies*

In the technology adoption literature it is standard to assume that decisions are based on a classical expected utility criterion using a von Neumann and Morgenstern approach. Underlying this concept is the principle that expected utility is computed in terms of a probability measure that accurately models future and unobserved trait evolutions. As argued above, however, the computation of this expectation measure might be itself subject to uncertainty (ambiguity). Below, I present the estimation strategies that correspond to three different decision rules and assumptions about the information structure and computational ability of the decision-maker.

The first strategy presumes that the decision maker can infer the true distribution of the outcomes or traits of interest and does not sort available information signals (because all signals are considered informative about the local nature of the uncertain process of interest). Thus, this strategy is more appropriate for scenarios with experimental conditions and agents with high computational ability. This will be identified by the name of *full observability case*. The second is consistent with agents whose observability is limited, who have limited computational ability or who sort their information because they consider that the environment faced by farmers located in different information regions is fundamentally different to their own. However, these agents must be confident about being able to do accurate predictions of outcomes under alternative scenarios (despite their potential reduced information or low computational ability). This will be identified by the name of *standard boundedly rational case.* The third strategy is as the second one except for the fact that it allows for the consideration of a full range of expected utility values associated with a complicated environment (defined in the introduction). This will be identified by the name of *ambiguity boundedly rational case.*

4.3.1 Full observability case: the farmer *i* knows the objective distributions of traits (conditional on his information set I_i), and therefore has an idea about the mean of this distribution, $E[X_j| I_i(t)]$, he solves the problem:

$$\text{Max}_j\ E[U_{ijt}| I_i(t)] \qquad i=1,\ldots,I; \quad j=1,\ldots,J \qquad (16a)$$

The researcher solves the problem

$$\text{Max}_j \int E[U_{ijt}| I_i(t)]\ dP(e) \qquad (16b)$$

When the farmer does not have full observability, but only knows that the vector of expected mean traits belongs to some set $R[E(X_j)]$, the problem above is not solvable in general.

4.3.2 Standard Bounded-Rationality Case: the farmer considers himself as a member of a reference group N_i that he predicts to have similar production conditions to his own. He then solves the *limited-information* version of the previous model:

$$\text{Max}_j\ E_N[U_{ijt}| I_i(t)] \qquad (17)$$

and the researcher solves the analogous problem as in (16b)

4.3.3 Ambiguity-Bounded-Rationality Case: a farmer might consider himself as a member of a reference group and still think that this reference group does not provide him with enough information as to know what he would obtain in his farm when growing a new variety. If this is the case then $R[E_N(X_j)]$ might contain multiple values. Then, for the researcher to solve the problem it is necessary to figure out how the farmer might choose among those values.

Manski (2004) has reviewed some "reasonable" decision rules in this setting. One is the *maximin* rule, proposed by Wald(1950). This rule is used by decision-makers who are ambiguity averse and prefer to make their calculations based on a bad scenario, rather than being optimistic and getting disappointed. In the present setting, this amounts to solving the problem:

$$\text{Max}_{j}\ \text{Min}_{\eta\in\ \text{R[EN}(Xj)]}\ E_N[U_{ijt}] \qquad (18)$$

The farmer thinks of the worst feasible state of the world according to the information in his reference group, and makes his choice maximizing his utility given that worst case scenario.

Another reasonable rule, proposed by Hurwicz (1951) consists of the maximization of a weighted average of the minimum and maximum values of the objective function that are feasible for each variety:

$$\text{Max}_{j}\ \Theta_{j}*\{\text{Min}_{\eta\in\ \text{R[EN}(Xj\mid Ii(\text{t}),\,\Delta i]}\ E_N[U_{ijt}]\}+(1-\Theta_j)*\{\text{Max}_{\eta\in\ \text{R[EN}(Xj\mid Ii(\text{t}),\,\Delta i]}\ E_N[U_{ijt}]\} \qquad (19)$$

$\Theta_j \in [0,1]$, expressing different degrees of optimism or pessimism.

Unlike the decision rule in equation 19, the one represented by equation 18 has been axiomatized by Gilboa and Schmeidler (1989) and shown to correspond to preferences that account for aversion against risk and ambiguity (see also Follmer and Schied 2004; Section 2.5). Taking the infimum of all expected utilities (or minimum if the set of expected utilities is bounded below) for prior models that describe the beliefs about possible future scenarios corresponds to a worst-case approach. This type of preference ordering is increasingly being used in analysis of optimal investment decisions (Talay and Zheng 2002, Quenez 2004, Schied 2004, 2005; Burgert and Ruschendorf 2005, Muller 2005, and Hernandez-Hernandez and Schied 2006).

An important feature of the *ambiguity-bounded-rationality* estimation strategy is that it is a generalized form of the *standard bounded rationality* one. In other words, for some farmers, equations 18 and 19 collapse to the decision rule described in equation 17. Recall that the degree of ambiguity that a farmer faces determines the size of the sets of expected traits and, consequently, of expected utilities. Since this degree varies across farmers --more specifically, it decreases as the amount of information signals received by a farmer as well as his computational ability increase, or as the local conditions are more homogeneous— for some farmers $\psi_j(.)$

tends to one and this implies that $E_N[U_{ijt}] = E_N[U_{ijt}| I_i(t)]$ = Min $_{\eta\in R[EN(Xj)]}$ $E_N[U_{ijt}]$ = Max $_{\eta\in R[EN(Xj)]}$.[19] Thus, the *standard bounded rationality* estimation strategy is a special case of the latter strategy, where all farmers have high confidence in the data observed and predictions made, perhaps because their farming conditions are highly homogeneous.

5. Discussion: Implications of the Different Frameworks

Implementing the model through a Bayesian mechanism presupposes having all the information that is necessary to infer all past distributions of outcomes, knowing the true noise of information and having excellent computational ability. These assumptions result in an expected utility of the type that is valid in a *full observability* setting --identified with equation 6'-- for all farmers, which corresponds to one special case of $R\{E_N [U_{ijt}]\}$ being a singleton, and where farmers do not need to sort available information. Thus, it enables choice through the standard utility maximization method described by equation 16a (in Decision Rules section), because the objective function is uniquely identified and expectations are not conditional on the information that the farmer belongs to an information region.

In this framework, a marginal change in the number of observations over time and across sources unambiguously reduces the prediction error (which can be established from equation 11). This is a consequence of the assumption that all information pieces are independent and reduce the prediction error by the same amount and regardless of the time when they were observed. Thus, as the number of trials observed becomes larger, the uncertainty about X_j tends to zero (then, it does not seem surprising that 'under full observability' X_j becomes 'fully observable'). The decrease in variance of the prediction error caused by trial observations, in turn, reduces the weight assigned to new information. Intuitively, as the technology becomes more popular and the experience of the farmer with the new variety increases, there are less unexpected elements and the weight that farmers assign to this type of information decreases. These implications seem to be more realistic for a technology that is, *ex ante*, known to bring a higher average utility to the farmer population considered. Several studies actually assume the superiority of the new

[19] Complete homogeneity of farming conditions is a very unlikely scenario and thus is not considered in detail. It would imply that $x^N_{j0}= x^N_{j1}$ in equation 7.; that is, the set of feasible outcomes is single valued; implying that Min $_{\eta\in R[EN(Xj)]}$ $E_N[U_{ijt}]$ = Max $_{\eta\in R[EN(Xj)]}$. Although it does not directly imply $E_N[U_{ijt}]= E_N[U_{ijt}| I_i(t)]$, I assume that for farmers with such a feasibility set either $E_N[X_j | I_i(t)] = x^N_{j0}= x^N_{j1}$ or $\pi_j(.)=1$. In both cases, $E_N[U_{ijt}]= E_N[U_{ijt}| I_i(t)]$ holds.

variety as a technological characteristic (Foster and Rosenzweig, Isik and Khanna 2003). However, if a new technology is characterized by its traits and these are perceived differently by each individual, the average level of utility resulting from the adoption of this technology is itself an empirical question. In fact it is often the relevant question in the case of new technologies.

In this standard approach, the stationarity characteristic of the trait distribution is what enables the farmer to *learn* about his own trait distributions by observing past experiences. [20] Thus, not only can learning occur very fast, but it is easy (because mean and variance of the process are constant over time and across locations). Even if the farmer is assumed to receive information signals from a limited group of people (reference group or neighbors), most applications using Bayesian learning assume that all farmers care about knowing the same average trait level (X_j) rather than about a local level (X_j^N). This assumption, as explained earlier, is related to the underlying stationarity of the process. If this unique mean for a variety j is higher than the mean for another variety k, given that the prediction error of farmers reduces over time as more information is available (although potentially enhancing optimization performance for the econometrician), this results in a quick loss of diversity in the chosen alternatives in the population and generalized adoption of the highest mean payoff variety.

As a result, lack of adoption, transfer or disadoption of new technologies is hardly explained under the assumptions of the model. And, depending on the type of information collected, this approach can create selection bias, if the analysis relies only on adopters; and specification bias, if the study deals with individuals who have used the technology some time but the model does not allow for disadoption. In fact, many studies that have adopted the Bayesian framework seeking to explain adoption in a country have used samples including only in the regions where adopters are to be found (Besley and Case, Foster and Rosenzweig, Cameron 1998, Munshi 2004). Another reason related is that the model has little ability to explain why a farmer who has no experience with a new variety, and who has no neighbors growing this variety, would adopt it. In fact, the expected traits of new technologies for individuals living in regions where little or no adoption has yet taken place are almost arbitrarily determined by the initial belief structure assumed. In many applications this initial condition is just constant across individuals or takes the form of a uniform –uninformative- prior. And including these individuals as part of the study results in a loss of explanatory power, at best, or

[20] See Manski 2004 for a different framework using the same assumptions and with similar implications.

in a more serious estimation problem if the number of adopters is very small (e.g. in the initial phases of introduction of a technology). This problem is analogous to the problem of finding an adequate ’reservation utility’ level for non-adopters, and has been addressed by labor economists through search models and by sociologists through the concept of heterogeneous “adoption thresholds”. [21]

In contrast to the standard Bayesian assumptions, the alternative approach assumes that farmers cannot observe the full distribution of traits and they think that it is difficult to infer it from the observed data. Their lack of confidence on the observed data or on their ability to process it and their uncertainty about a measure of expected utility associated with unobservable factors leads them to use a choice criterion that differs from standard utility maximization (Equation 18). However, the degree of confidence differs across agents and the choice criterion resembles the standard utility maximization for farmers with high confidence (Result 1.a). In this case, the choice criterion in equation 18 collapses to the one in equation 17.

Because of the farmers’ perception that it is difficult to infer the full trait distribution, they try to approximate the mean of the local distribution of traits by sorting out the information received. Since the full structure of information noise is not known by farmers, they sort out their data relying on a representativeness or similarity criterion, in each specific period. While standard frameworks see the neighborhood size as an informational constraint that limits learning and adoption, the alternative approach also emphasizes the ability to sort and combine information (Results 4) as a limitation. It is farmers in “regions” with low adopter rates, who have not grown the variety under consideration recently (Result 6) and whose location is closest to the average adopter location who benefit the most from information access (because observed information is most precise for them). Farmers who have no own experience growing a new variety might see their confidence increased because they see more people growing the variety in a specific period. If they are ambiguity averse, this might lead them to adopt the variety. However, it is the own experience in the next period which will determine whether the farmer will continue growing the variety or disadopt. In general, choices by farmers are influenced by the popularity of a variety under consideration, as in the standard framework; however, here they are also influenced by the popularity of similar varieties. As shown in chapter 1, a trait-based

[21] In the model above, these thresholds vary across individuals according to the performance of traditional varieties in their location.

framework facilitates the analysis of multiple related varieties and as derived in this chapter, it captures the possibility of disadoption or transfer to a similar variety (Result 2).

Computational ability is different across farmers and it influences negatively the degree of unmeasurable uncertainty that a farmer faces. Thus, for farmers with a constant over time low computational ability, the prediction uncertainty (defined in this setting as the difference between mean expectation and true mean trait level in the location) might never be zero, even if they observe an infinite number of observations. In fact, as the confidence of a farmer increases, the ambiguity faced by him decreases, but this does not imply that his beliefs converge to the true local mean of the process. In terms of the model, a farmer's approximation of the mean trait, $E_N[X_j]$, does not necessarily converge to the true X_j^N. If this convergence implies learning, then learning in the model depends on whether the process is stationary or not.

To clarify things, there are two main types of convergence here (for a given location). One is the one just mentioned ($E_N[X_j]$ to the true X_j^N), and which implies learning over time about the true local mean outcome, by farmers in a specific location. The second is the convergence of $E_N\ [X_j|\ I_i(\text{t})]$ towards $E_N[X_j]$ and represents maximum reduction of ambiguity over time and across farmers.

The first type of convergence – learning - in a scenario with ambiguity, can be seen as a sequential reduction of this type of uncertainty, as Manski (2004) has shown. He proposed a model of social learning under ambiguity, showing that assuming stationarity of the process of interest, ambiguity would lead to lack of convergence to the true mean. In particular, if agents are pessimistic, the adoption rate of an innovation increases with time and converges to a steady state that is below the optimal adoption rate. On the contrary, if agents are optimistic, the adoption rate starts at a high level, decreases with time and falls to a steady state that is above the optimal rate of adoption. However, Manski does not consider learning-by-doing in his model, which is a highly unbiased source of information for agricultural scenarios. The effect of this type of information will likely cause convergence to occur closer to the true mean than in Manski's scenario. However, achieving a steady state highly depends on the stationarity assumption.

In highly non-stationary environments learning is more difficult because -- as Suppes (1994) puts it when referring to non-stationary environments-- "in the worst kind of environment no significant event has a high probability". Thus, it is difficult to identify any regularity in outcomes over time in order to make any accurate prediction. One can think of stationarity as increasing information representativeness in an intertemporal sense; when the trait process is stationary, data originated in every period is representative of the whole trait process and the best estimator is an average over time of the stored data.. If the process is non-stationary, a farmer might be best considering only recent information. In this case, traits are perceived as changing constantly, and only agents who have an a priori knowledge of the way in which these changes vary can *learn* about the traits of a technology by tracking the data generating process over time (As noted by Epstein 2006 (P.19), this is what happens in regime switching models with persistent hidden state variables). However, learning never ceases because agents are always at least one step behind.

Convergence of expectations across locations depends on heterogeneity as well. Homogeneity across locations means that there are less and wider information regions and that information is more representative of every individual farmer randomly selected from the population. Thus it amounts to having more information. High heterogeneity across locations implies more fractured information regions and slower convergence.

For ambiguity-averse individuals, reductions in ambiguity or unmeasurable uncertainty with respect to one variety increase the adoption probability. As results 1a) and b) imply, ambiguity decreases as computational ability, own experimentation and popularity increase, and is lower for individuals living in more homogeneous regions (Homogeneity of farming conditions not only reduces the local spread of outcomes but also translates into more information representativeness and more similarity of the varieties grown and, therefore, in increased available information for individual farmers and consequent ambiguity reductions). However, if unobservable conditions and complexity persist over time, farmers might never be completely confident that the observed data will yield good predictions (the process is never perceived as being stationary and signals are never perceived as being iid). Thus, at the population level there might be a point where no further ambiguity reduction is possible even if access to information is complete. Although information costs are not directly incorporated in the model, costly acquisition of computational ability would translate into lower computational ability and a higher degree of ambiguity.

6. Concluding Remarks

Key assumptions maintained in technology adoption studies using standard Bayesian learning mechanisms are found to be more appropriate for experimental settings than for agriculture. These identified assumptions are: First, the true distribution of relevant variables such as traits and information signals is assumed to be fully observable. Second, information signals (about trials) are independently and identically distributed (and so are they perceived by farmers). The noise of these signals is assumed to be common to and known by the farmers. Thus, farmers can be (and are) sure about the connection between these signals and the true parameters of the fully observable trait distribution. Third, decision-makers have infinite memory, enormous computational capacity (they use large sets of trial data and complicated optimization algorithms to make decisions) and have great ambitions about what they can learn and observe (they have no doubt that beliefs converge to the truth). Fourth, the unknown factors to be learned about (parameters of the trait distribution) do not change over time or across locations (learning is easy!). These assumptions determine that farmers get to completely characterize the true mean traits of a technology over time.

However, there exist several reasons why agents choosing among various agricultural technologies are not likely to fully observe trait distributions or rely on data intensive and mathematically complicated optimization algorithms; implying that learning (from past experiences and from others) is more difficult in agricultural settings than it would be in laboratory type scenarios. Therefore, standard approaches create biases in estimations because of their informational assumptions. In particular, one main source of measurement bias arises because the farmer does not estimate a mean trait based on all observed signals, but rather cares about the quality of these data (in terms of the information that they convey regarding potential outcomes on his own farm) and sorts out the information he receives before processing it.

A second source of measurement bias is related to the perception of the farmer that the data he has observed in his location are not a random sample. That is, that information signals are not based on trials that are representative of the whole population in his information region. The bias related to this source increases as farmers' computational ability and experience with closely related varieties decreases. It is also high for farmers whose farming conditions differ mostly from the conditions of early technology adopters.

A technology adoption model that neither relies on full observability of trait probability distributions nor *iid*-ness of information signals has been proposed. The required assumptions are that farmers use heuristic rules to decide between data sources and a notion of data representativeness to sort information. Farmers sorting of information separates the population in heterogeneous groups that differ in their information structures and defines different adoption determinants across groups. The model allows for farmers to be unsure about the reliability of data and about their ability to make predictions based on these data, leading them to not only sort information signals but to potentially face a type of uncertainty that is termed ambiguity. The complex nature of the environment, the lack of confidence and the uncertainty faced by farmers influences their behavior and might lead them to never cease learning.

Two alternative estimation approaches to the standard rational expectations optimization framework are proposed. One is based on standard bounded-rationality considerations and involves maximizing a local expectation determined by the information region of the farmer. The second is grounded on a combination of ambiguity and bounded-rationality, which represents a generalization of the former. These estimation strategies will be applied in the next chapter of this dissertation. The econometric model used for this application will be the genetically-modified trait based adoption model developed in the first chapter of the dissertation. This will provide a means to see how the different strategies perform, and how sensitive the estimation results in chapter one are to alternative formulations of the problem that were developed in this second chapter.

Finally, the model developed in this article is not seen to be relevant only for applications related to adoption of innovations as the one presented. It relates to any setting were decision-makers might be uncertain about the consequences of their acts or attributes of their choice alternatives, and do not feel confident in assessing probability distributions of these consequences or attributes. As an example, studies seeking to perform hypothetical valuation of goods can be seen as sharing such a setting. These studies are widely criticized because of the potential uncertainty faced by respondents when they are asked to make choices. Studies which have dealt with this uncertainty take the view that some respondents are unsure what their behavior would be in an actual scenario (See Moore 2006, referring to willingness-to-pay behavior). This article proposes that the uncertainty is not related to uncertainty regarding how to behave, but rather, is due to lack of or unobservability of information determining the true obtained traits in the actual scenario.

Bibliography

Bandiera, O. and I. Rasul. 2005. "Social Networks and the Adoption of New Technology in Northern Mozambique" CEPR Discussion Paper no. 3341. forthcoming Economic Journal.

Besley, T. and A. Case. 1994. "Diffusion as a Learning Process: evidence from HYV cotton." Princeton University. Woodrow Wilson School of Public and International Affairs. Research Program in Development Studies. Discussion Paper No. 174: 1-19

Besley, T., and A. Case. 1993. "Modeling Technology Adoption in Developing Countries." *The American Economic Review* 83(2): 396-402.

Burgert, C. and L. R ¨ uschendorf. 2005. "Optimal Consumption Strategies Under Model Uncertainty." *Statist. Decisions* 23 (1): 1–14.

Cameron, L. A. 1999. "The Importance of Learning in the Adoption of High-Yielding Variety Seeds," *American Journal of Agricultural Economics* 81: 83-94.

Conley, T. and C. Udry. 2001. "Social learning through networks: the adoption of new agricultural technologies in Ghana." *American Journal of Agricultural Economics* 83(3): 668-73.

Debreu, G. (1959) *Theory of Valu*e, Wiley: New York. Ellison, G. and D. Fudenberg. 1993. "Rules of Thumb for Social Learning." *Journal of Political Economy*, 101(4):561-758.

Ellsberg, D. 1961. "Risk, Ambiguity and the Savage Axiom." *Quarterly Journal of Economics* 75, 643-669.

Epstein, L. 2006. "An axiomatic model of non-Bayesian updating," *Review of Economic Studies* 73:413-436.

Epstein, L.G. and M. Schneider. 2006 ."Learning under ambiguity," Working Paper, Department of Economics, University of Rochester, Rochester: NY.

Epstein, L.G. and M. Schneider. 2002. "Recursive multiple-priors." *Journal of Economic Theory.*

Epstein, L.G. and J. Zhang. 2001. "Subjective probabilities on subjectively unambiguous events." *Econometrica* 69, 265-306.

Epstein, L.G. and T. Wang. 1994. "Intertemporal asset pricing under Knightian uncertainty." *Econometrica* 62, 283-322.

Gilboa, I. and D. Schmeidler. 1989. "Maxmin Expected Utility with Non-Unique Priors." *Journal of Mathematical Economics* 18, 141-153.

Feder, G. and G. T. O'Mara (1982). "On Information and Innovation Di¤usion: A Bayesian Approach." *American Journal of Agricultural Economics* 64(1): 145-47

Feder, G., and D. Umali. 1993. "The Adoption of Agricultural Innovations: A review." *Technological Forecasting and Social Change*, 43: 215-239.

Follmer, H., Schied, A. 2004. *Stochastic Finance: An Introduction in Discrete Time* (de Gruyter Studies in Mathematics 27). Berlin: Walter de Gruyter 2002. Second revised and extended edition.

Foster, A. and M. R. Rosenzweig. 1995. "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." Journal of Political Economy 103:1176-1209.

Hernandez-Hernandez D., and A. Schied. 2006. "Robust Utility Maximization in a Stochastic Factor Model." *Stat. Decisions* 24, 109-125 (2006).

Hurwicz, L. 1951. "Optimality criteria for decision making under ignorance," *Cowles Commission Discussion Paper* 370, Chicago.

Isik, M. & Khanna, M. 2003. Stochastic technology, risk preferences, and adoption of site-specific technologies. *American Journal of Agricultural Economics*, 85(2):305-317.

Knight, F. 1921. "Risk, Uncertainty, and Profit." Boston: Houghton Mifflin.

Krantz, D. H., R. D. Luce, , P. Suppes & A. Tversky. (1971) *Foundations of Measurement, Vol. I: Additive and Polynomial Representations*. New York: Academic Press, 577 pp.

Manski, C. 2005. *Social Choice with Partial Knowledge of Treatment Response*. Princeton University Press: Princeton NJ.

Manski, C. 2004. "Social Learning from Private Experiences: The Dynamics of the Selection Problem," *Review of Economic Studie*s, 71, 443-458.

Manski, C. 2003a. "Measuring Expectations," *Econometric*a, 72, 1329-1376.

Manski, C. 2003b. *Partial Identification of Probability Distributions.* New York: Springer-Verlag .

Manski, C. 1995. *Identification Problems in the Social Sciences.* Cambridge: Harvard University Press.

Manski, C. 1993. "Adolescent Econometricians: How Do Youth Infer the Returns to Schooling?" in C.Clotfelter and M. Rothschild (editors), *Studies of Supply and Demand in Higher Educatio*n, Chicago:University of Chicago Press, 43-57.

Marra, M.C., B. Hubbell, G.A. Carlson. 2001. "Information Quality, Technology Depreciation, and Bt Cotton Adoption in the Southeast." *Journal of Agricultural and Resource Economics* 26(1):158-75

McFadden, Daniel, 1999. "Rationality for Economists?," *Journal of Risk and Uncertainty*, Springer, vol. 19(1-3):73-105, December.

Munshi, K. and J. Myaux. 1998. "Social Effects in the Demographic Transition: Evidence from Matlab," Bangladesh. Mimeo. University of Pennsylvania.

Munshi, K. 2002. "Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution." Mimeo. University of Pennsylvania.

Quenez, M.C. 2004. Optimal Portfolio in a Multiple-Priors Model. Seminar on Stochastic Analysis, Random Fields and Applications IV, 291–321, Progr. Probab., 58, Birkhäuser, Basel.

Shackle, G.L.S. 1955. "Uncertainty in Economics: And other reflections." Cambridge, UK: Cambridge University Press.

Schmeidler, D. 1989. "Subjective Probability and Expected Utility Without Additivity." *Econometrica* 57, 571-587.

Shafir, E. (Ed). 2004. *Preference, Belief, and Similarity*: Selected Writings of Amos Tversky. Cambridge, MA: MIT Press.

Simon, H. 1959. "Theories of Decision-Making in Economics and Behavioral Science", *American Economic Review* 49, 253-283.

Simon, H.A. 1955. "A behavioral model of rational choice." *Quarterly Journal of Economics*, 69, 99-118.

Slovic, P., Kahnemann, D., and Tversky, A. (eds.) 1982. *Judgement Under Uncertainty: Heuristics and Biases.* Cambridge: Cambridge University Press.

Suppes, P. 1994. Learning and Projectibility. In D. Stalker (Ed.), *Grue! The new riddle of induction*. Chicago, IL, Open Court, 263-272.

Suppes, P. 1994. Qualitative Theory of Subjective Probability. In G. Wright and P. Ayton (Eds.), *Subjective Probability*. John Wiley & Sons, 17-37.

Suppes, P. 1986. Comment on "The Axioms of Subjective Probability," by Peter C. Fishburn. *Statistical Science*, **1**, 347-350.

Schied, A. 2005. "Optimal Investments for Robust Utility Functionals in Complete Market Models." Math. Oper. Research. 30, No. 3, 750-764.

Schied, A. 2006. "Risk Measures and Robust Optimization Problems." Stochastic Models, 22,753-831.

Schied, A., C.T. Wu. 2005. "Duality Theory for Optimal Investments Under Model Uncertainty." Stat. Decisions 23, No. 3, 199-217.

Talay , D., and Z. Zheng. 2002. "Worst Case Model Risk Management." Finance Stochast. 6, 517-537.

Tversky, A. and Kahneman, D. 1973. "Availability: A heuristic for judging frequency and probability." *Cognitive Psychology*, 5, 207-232.

Wald, A. (1945), "Statistical Decision Functions which Minimize the Maximum Risk," *Annals of Mathematic*s, 46: 265-280.

Wells, G.L. & Harvey, J.H. 1977. "Do People Use Consensus Information in Making Causal Attributions?" *Journal of Personality and SocialPsychology*, 35, 270-293.