Sequential Adoption of Package Technologies
The Dynamics of Stacked Trait Corn Adoption

Ursula Aldana*, Jeremy D. Foltz*, Bradford L. Barham*, and Pilar Useche**
***University of Wisconsin – Madison, **University of Florida**

Contact Author: Brad Barham
barham@mailplus.wisc.edu
608-265-3090

Mailing Address:
Dept of Ag and Applied Economics
427 Lorch Street
Madison, WI 53706

January 2010

Abstract:

GM corn seed companies have innovated continuously with the introduction of new traits and, more recently, with the creation of stacked varieties, which combine more than one trait. This work develops a Bayesian model of adoption dynamics that demonstrates how uncertainty with a package technology with known risk can lead to a sequential adoption pattern in which farmers adopt a single component first. We then develop a semi-parametric panel data model of adoption dynamics to measure the effects of experience with single trait (non-stacked) varieties on the adoption of stacked varieties. The results underscore the importance of early experience with the non-stacked technology in the subsequent adoption of stacked varieties, i.e., a sequential adoption process. There is also evidence that farmers with more human capital tend to learn faster from own experience and that as the GM corn-technology diffusion process deepens, the importance of early experience decreases.

Acknowledgements: This work was funded through the Program on Agricultural Technology Studies core budget, a UW-Hatch grant and the Illinois Missouri Biotechnology Alliance grant from the USDA. This work has benefited from discussions with Gustavo de los Campos, Michael Carter, Jean-Paul Chavas, Paul Mitchell, Terry Hurley, Laura Schechter, and participants at the Applied Economics Workshop at University of Wisconsin-Madison.

Sequential Adoption of Package Technologies
The Dynamics of Stacked Trait Corn Adoption

Most recent models of technology adoption take the new technology to be a single uniform technology that firms or farmers decide whether to use. Many technologies in fact consist of sub-components that are available individually or jointly as a package. Genetically modified (GM) corn to date presents such a technology option to farmers, with single trait or multiple (stacked) trait varieties. This divisibility of the technology package can create a particular adoption dynamic, since the experience gained through the adoption of subcomponents will provide information on the characteristics of the package thereby encouraging or discouraging its subsequent adoption.

A now old literature on green revolution technology adoption from the 1970's and 1980's analyzed package technologies, and identified sequential adoption patterns in which farmers adopted parts of the package before adopting the whole package (see e.g., Byerlee and Hesse de Polanco; Leathers and Smale). These works were primarily empirical, and argued that fixed costs, credit constraints, risk, uncertainty, and learning all contributed to a sequential adoption pattern. In the US GM corn market, we posit that learning and uncertainty are likely to be the most salient of these explanations. GM seeds are inherently divisible, US farmers are not likely to face credit constraints that bind the purchase of a sufficient quantity of seeds to test GM varieties, and crop insurance is available to help manage basic risks. [1]

[1] During the period of our data, 2000-2006, US grain farmers had historically easy access to credit at low interest rates. In addition, most farmers in our sample purchased crop insurance. Thus, while credit constraints and risk aversion may play a role in the sequencing of adoption, they are likely to be a small contributor to the issue.

Technology adoption has long been viewed as a process that hinges on farmer learning about uncertain and risky options (Griliches; Rogers; Feder and Slade; Feder et al.; Foster and Rosenzweig; Conley and Udry). Often the learning process about a new technology can drive the timing of adoption. As explained in Feder and Omara, uncertainty diminishes over time because of the common experience gained through the adoption of farmers in the economy and because of the expected improvement of extension services and advancement of research studies. Similarly, uncertainty decreases with direct experience with the technology or with a component of the technological package. A strand of the literature has modeled this decrease in uncertainty using a Bayesian approach (Lindner; Feder and Slade; Leathers and Smale), according to which producers update their beliefs about the distribution of the profitability of the technology, using observations on the profit achieved by other producers or themselves. From this literature, the Bayesian conceptual model by Leathers and Smale is the one that comes closest to our work in providing a reason for a sequential adoption process. Our work expands the Bayesian adoption literature by considering correlated learning across components of package technologies and in so doing provides a new logic for a sequential pattern of adoption.

Previous studies of farmer adoption of GM crops focus almost exclusively on the binary decision to use GM varieties (Alexander, Fernadez Cornejo et.al., Hubbell et.al.) or else on a one-time decision across multiple independent or correlated choices (Useche et al.). All of these GM adoption studies are cross-sectional analyses that preceded or coincided the introduction of stacked GM traits, and do not account for timing or previous experience in a systematic fashion. Thus, none of these studies explicitly account for the potential sequencing of farmer adoption choices, especially the potential path of trying one or more single variety traits before adopting stacked varieties, which combine multiple traits. Examining the adoption of stacked varieties through a dynamic lens with longitudinal data can improve our understanding of the sequencing of farmer choices in this and other types of evolving package technologies.

In this article, we analyze the sequential nature of the adoption of stacked GM corn varieties, which are a package of individual GM traits. The main empirical questions that we address are whether and why farmers might move sequentially from a single trait to a stacked variety. In the GM corn market during our study period there were three traits (herbicide tolerance (HT), bacillus thuringeiensis (BT) to kill corn borers, and to kill corn rootworm insects) in the market that were sold both independently and as a packages. We develop a novel theoretical framework based on the concept of Bayesian updating of the beliefs regarding the expected profits of the stacked varieties. Using theory and data from a survey of corn farmers from Minnesota and Wisconsin conducted in 2006, we show that the possibility of learning about the profitability of the stacked variety through the use of a single trait variety reduces the uncertainty of stacked varieties, which tends to favor a sequential pattern of adoption.

A Bayesian Model of Adoption Dynamics

In this section, we develop a novel multivariate Bayesian model that explains the logic for a sequential adoption pattern for a package technology with individual components that could provide information on the value of the package. While this builds on a standard Bayesian learning framework (e.g., Anderson, Stoneman), it is to our knowledge the first model using a multivariate Bayesian framework to understand how farmers learn about the profitability of package technologies.

Our model features a seed market where the farmer chooses between conventional, single GM trait, and a stack of multiple varieties. In this Bayesian model, farmers learn about the yields of the stacked variety by using a single trait variety. Farmers make choices for the current period based on their current information set, which includes past experience with a GM trait. This allows us to analyze how experience with a single GM trait changes the incentives to adopt the package given the information set available to the farmer.[2]

Farmer's maximization problem

We assume that farmers face a choice set that includes J types of corn seeds (e.g., Conventional (c), HT only (1), BT only (2), and stacked HT/BT (s)). To simplify the modeling, we focus on the case in which the farmer adopts trait 1, and how that might create learning about trait 2 and stacked traits. The farmer will choose the variety j*, which gives her the highest expected utility:

$$j^* = \arg\max_j \{EU_j\} \tag{1}$$

where EU_j is the expected utility achieved when adopting variety j[3]. Following Stoneman, we express the farmer's expected utility as a linear function of the expected value as well as the variance of profits[4]:

---

[2] We do not explicitly model the potential forward-looking dynamics of farmer choices, the case where farmers might adopt a particular variety today because of the future value of information it might provide. This would unnecessarily complicate the modeling exercise and merely reinforce the sequential adoption logic we are seeking to demonstrate.
[3] In this model we constrain farmers to choose one variety at a time. An alternative specification would have the farmer choose a land portfolio in order to maximize her expected utility. Specification (1) takes into account risk management through the binary

$$EU_j = E(\pi_j) - bV(\pi_j) \qquad (2)$$

where b represents the farmers risk aversion level. As detailed below, we assume that the farmer does not know the expected profits of GM varieties, and this implies that the variance and the expected value in (2) represent both the parameters of the real distribution of profits and the parameters of the farmer's beliefs. The term $E(\pi_j)$ represents the mean of the distribution of the farmer's beliefs regarding expected profits, which may or may not be accurate to the true expected profits. The aggregate variance, $V(\pi_j)$, is the sum of two elements: the variance of the distribution of the beliefs regarding expected profits, and the variance of these profits implied by variation in common known but stochastic factors such as weather conditions. The variance of the beliefs can be interpreted as the uncertainty component while the variance of the profits implied by weather conditions represents what is typically thought of as production risk.

Profits depend on the output price, the yield, the input cost as well as the price of a particular seed variety in the following manner:

$$\pi_j = p^m * y_j - w_j - p_j^s \qquad (3)$$

---

adoption decision but abstracts from risk management through land portfolio allocation. Specifying the problem without the portfolio selection issue, however, allows us to analyze the impact of experience in a simpler setting. As shown in Feder and O'Mara, under the alternative portfolio selection specification farmers' decision to adopt a new technology (stacked seeds in this case) will depend on the difference between the expected utility from two land portfolios (one where land under the new technology is zero and another where this amount is the optimal one) and on the amount of fixed costs of adoption. Changes in the expected utility of a particular seed, as expressed in (2), will be positively correlated with changes in the expected utility of the portfolio that includes that seed. In that sense, the results obtain here will resemble those obtained under this alternative and more complicated specification.

[4] This standard mean-variance functional form is consistent with a farmer who has a CARA utility function when the actual profits and the farmer beliefs regarding the mean profits are normally distributed.

where p^m is the output price, y_j is the yield achieved by variety j, w_j represents the input costs incurred when variety *j* is used and p_j^s is the price of seed of variety j. We assume the yields, as well as the input costs are stochastic and depend on the variety chosen. Input costs are distributed normally, with a mean value of μ_{w_j} and a variance of $\sigma_{w_j}^2$, both of which we assume are known by the farmers. The distribution of yields is described below.

The yields achieved for each type of variety depend on the traits included in a particular seed. Define I_{jk} as an indicator function that equals one if variety *j* contains trait k, and define I_c as an indicator function that equals one if variety *j* is the conventional variety. Assuming that the number of traits is equal to 2, the yields can be expressed as:

$$y_j = \delta_c I_c + \delta_1 I_{j1} + \delta_2 I_{j2} + \delta_{12} I_{j1} I_{j2} \tag{4}$$

Equation (4) indicates that the yield of a conventional variety is equal to δ_c, while the yields of GM varieties depend on the traits contained in the seeds. In this formulation δ_{12} captures the impact on yields of having more than one trait present in a seed variety. We assume that δ_c, δ_1 and δ_2 are random variables which are distributed as $N(\mu_i, \sigma_i^2)$, where i=c,1,2; while δ_{12} is a random variable distributed as $N(\mu_{12}, \sigma_{12}^2)$; and that these variables are independent from each other. Note that the parameters $\sigma_c^2, \sigma_1^2, \sigma_2^2, \sigma_{12}^2$ represent production risk.

Farmer Learning about the Technology

In order to capture the uncertainty in the profitability of a new technology, we assume that the farmer knows the expected value of the yield of the conventional variety, μ_c, the variance of all yields ($\sigma_c^2, \sigma_1^2, \sigma_2^2, \sigma_{12}^2$), but does not know the expected values of the yields of the GM varieties: μ_1, μ_2 and μ_{12}. Instead, the farmers have a prior distribution, which captures their expectations regarding the value of these parameters. This prior distribution is updated in a Bayesian manner into a posterior distribution, as the farmer experiments with the technology.

We conceptualize the package as having two correlated elements [μ_1, μ_2] that can be learned about through experience with any of these elements and then a third part, μ_{12}, which is the part of the package that can only be learned about through use of the package.[5] In this conceptualization, we assume that the prior distribution of [μ_1, μ_2] is bivariate normal with a positive covariance between the two varieties. This captures the idea that experience with one trait provides information on the qualities of the other trait. For example, if the same company produces both traits, the farmer is likely to expect better (worse) results for trait 2 when the results for trait 1 turn out well (badly).

Note that it is possible for the prior distribution of [μ_1, μ_2] to have a positive covariance while δ_1 and δ_2 are independent from each other. The assumption of a positive covariance between μ_1 and μ_2 means that the farmer believes that if the expected value of trait 1 yields is high, then the expected value of trait 2 yields is more likely to be high. The assumption of independence between δ_1 and δ_2 establishes that, regardless of the expected value of the distribution of [δ_1, δ_2], a good year for trait one does not imply that

[5] Our assumption that μ_{12} is uncorrelated with the other elements of the package is a simplifying assumption. The sequential learning results would only be reinforced if we allowed μ_{12} to be correlated with μ_1 and μ_2 (i.e., learning about all elements of the package).

a good year for trait 2 is more likely to happen. On the other hand, the prior distribution of μ_{12} is assumed to have a mean given by θ_{12}^0 and a variance given by V_{12}^0. In order to simplify the analysis, we assume that beliefs regarding μ_{12} are uncorrelated with beliefs regarding [μ_1 , μ_2].

With the Bayesian formulation, the prior distribution of [μ_1 , μ_2] is denoted as $p^o(\mu_1, \mu_2)$ and is described as:

$$p^o(\mu_1, \mu_2) = N\left[\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}\middle|\begin{pmatrix}\theta_1^0\\ \theta_2^0\end{pmatrix}, \mathbf{\Sigma}\right] \tag{5}$$

where [θ_1^0, θ_2^0] are the expected value of the distribution and Σ is the variance covariance matrix which equals:

$$\Sigma = \begin{pmatrix}V_1^0 & C_{12}^0\\ C_{12}^0 & V_2^0\end{pmatrix} \tag{6}$$

In order to illustrate the sequencing of adoption we assume that the farmer starts by growing a single trait variety that contains trait 1. She uses *n* observations (e.g., years growing the crop) on the yields of a single trait variety containing trait 1 to update her beliefs regarding μ_1. Given the assumption of a positive covariance between μ_1 and μ_2 in the prior distribution, these observations on trait 1 will also allow the farmer to update her beliefs regarding μ_2. Since both traits 1 and 2 are present in the stacked variety, observations on trait 1 will allow the farmer to update her beliefs about the mean yields of the stacked variety.

The learning process that uses observations on trait one to update the beliefs about this same trait can be conceptualized as a standard Bayesian problem in which the prior is a normal univariate distribution. Consequently, the posterior distribution of μ_1 will be distributed normally. Denoting this posterior as $p^1(\mu_1)$, we have that:

$$p^1(\mu_1) \sim N(\theta_1^1, V_1^1) \tag{7}$$

where the expected value θ_1^1 and variance V_1^1 can be calculated by the standard formula that corresponds to a Bayesian updating of the beliefs when the prior is a univariate normal distribution (Anderson). That is:

$$\theta_1^1 = \frac{\dfrac{\theta_1^0}{V_1^0} + \dfrac{n\overline{\delta_1}}{\sigma_1^2}}{\dfrac{1}{V_1^0} + \dfrac{n}{\sigma_1^2}} \tag{8}$$

$$V_1^1 = \frac{1}{\dfrac{1}{V_1^0} + \dfrac{n}{\sigma_1^2}} \tag{9}$$

In equation (8) the farmer's expected value of the posterior distribution will be a function of the expected value of the prior distribution, θ_1^o, the variance of that distribution, V_1^o, the known variance of yields, σ_1^2, her observed mean yield, $\overline{\delta}_1$, and the number of trials the farmer has had with the technology, *n*. We can see that if the average yield observed through experience is higher than the expected value of the prior distribution, the expected value of the beliefs regarding μ_1 will be updated positively. If the contrary is true, the expected value of the beliefs will be updated negatively. Note that, although not explicitly modeled, the mean of the prior distribution may be influenced by common knowledge of the technology outside of actual experience. Equation (9) shows the classical result that as the number of observations, *n*, increase, the variance of the distribution of the beliefs decrease, i.e., $\frac{\partial V_1^1}{\partial n} \leq 0$ and that $V_1^1 \leq V_1^o$. This variance can be regarded as the uncertainty faced by the farmer, an uncertainty that decreases as she learns through experience.

The posterior distribution of μ_2 can be deduced from the joint prior distribution, equation (5), together with the posterior distribution of μ_1, equation (7), in the following manner:

$$p^1(\mu_2) \sim \int p^o(\mu_2 \mid \mu_1) * p^1(\mu_1) d(\mu_1) = g(\theta_2^1, V_2^1) \qquad (10)$$

Where g is a distribution function with a expected value given by: θ_2^1 and a variance given by V_2^1. Taking into account the properties of the conditional distribution associated with a multivariate normal, we have that[6]

$$\theta_2^1 = \theta_2^0 + \frac{C_{12}^0}{V_1^0}(\theta_1^1 - \theta_1^0) \qquad (11)$$

$$V_2^1 = V_2^0 - \frac{(C_{12}^0)^2}{V_1^0} + \left(\frac{C_{12}^0}{V_1^0}\right)^2 * V_1^1 \qquad (12)$$

Definition 1: *A technology is "ascendant" when* $\theta_1^1 \geq \theta_1^0$ *and* $\partial\theta_1^1 / \partial n > 0$*, i.e., it is a technology for which new information has a non-negative effect on the mean of the distribution of the beliefs about the mean returns to the technology.*

Lemma 1: *When technology 1 is ascendant, and* $C_{12}>0$*, then* $\theta_2^1 \geq \theta_2^o$ *and* $\partial\theta_2^1 / \partial n > 0$

Lemma 2: *When* $C_{12} > 0$*, then* $V_2^1 \leq V_2^o$ and $\frac{\partial V_2^1}{\partial n} < 0$

Lemma 1 says that for correlated technologies, a positive experience with technology 1 will increase the farmer's mean of the distribution of the beliefs regarding the average yields of technology 2. Lemma 2 says that this experience with technology 1 will

[6] For a detailed proof see: Aldana et al.

also decrease the uncertainty with technology 2. Taking into account the definition of an ascendant technology and that θ_2^0 will satisfy equation (11) if we replace θ_1^1 by θ_1^0, Lemma 1 follows from equation (11). In the same manner taking into account that V_2^0 will satisfy equation (12) if we replace V_1^1 with V_1^0 and that $\frac{\partial V_1^1}{\partial n} \leq 0$, Lemma 2 follows from equation (12). This means that an increase in the expected value of the distribution of the beliefs regarding μ_1 implies an increase in the expected value of the distribution of the beliefs regarding μ_2 and that experience with trait 1 reduces the variance of the beliefs regarding μ_2 .

Lemma 3: *When n>0, then* $C_{12}^1 \leq C_{12}^o$ and $\frac{\partial C_{12}^1}{\partial n} < 0$

Proof: It can be shown that covariance of the beliefs will be equal to[7]:

$$C_{12}^1 = \frac{C_{12}^0 V_1^1}{V_1^0} \tag{13}$$

In this case we also have that C_{12}^0 satisfies expression (13) if we replace V_1^1 with V_1^0. Since V_1^1 is less than V_1^0, this means that the covariance of the posterior distribution will be lower than the covariance of the prior distribution. Since $\frac{\partial V_1^1}{\partial n} \leq 0$, the covariance of the posterior distribution will decrease as *n* increases. Lemma 3 shows that experience with trait 1 will also reduce the covariance of the bivariate distribution of the beliefs.

Farmer Choice of Technology:

Since producers choose the technology that provides the highest expected utility, equation (1), the changes in beliefs from trials with technology 1 will affect the technology choice

[7] For a detailed proof see Aldana et al.

through changes in the expected utility from adopting different technologies. Given that the farmer does not know the expected value of profits, but instead has a distribution of beliefs regarding its possible values, the expected utility she maximizes is given by:

$$EU_j = \overline{\pi}_j^{Bi} - bV_{ag}(\pi_j) \tag{14}$$

where $\overline{\pi}_j^{Bi}$ represents the farmer's best approximation to the expected profits, which is the mean of the distribution of the beliefs regarding expected profits. The superscript *i* is equal to zero if there is no experience with any genetically modified trait and equal to one if the farmer has some experience that allowed her to learn about the yields of the technology. In the second term of the right hand side of (14), $V_{ag}(\pi_j)$ represents the aggregate variance (described in detail below) including both risk and uncertainty. The expected value of the beliefs of profits will be:

$$\overline{\pi}_j^{Bi} = p^m * (\mu_c I_c + \theta_1^i I_{j1} + \theta_2^i I_{j2} + \theta_{12}^0 I_{j1} I_{j2}) - \mu_{w_j} - p_j^s \tag{15}$$

Specifying the crop choices, j as:

$$j = \begin{Bmatrix} 1 & \text{the seed only has trait 1} \\ 2 & \text{the seed only has trait 2} \\ s & \text{the seed has traits 1 and 2} \\ c & \text{the seed is conventional} \end{Bmatrix}$$

the expected value of the beliefs about profits for each of the technologies will be:

$$\overline{\pi}_c^{Bi} = p^m * (\mu_c) - \mu_{w_c} - p_c^s \tag{16}$$

$$\overline{\pi}_1^{Bi} = p^m * (\theta_1^i) - \mu_{w_1} - p_1^s \tag{17}$$

$$\overline{\pi}_2^{Bi} = p^m * (\theta_2^i) - \mu_{w_2} - p_2^s \tag{18}$$

$$\overline{\pi}_s^{Bi} = p^m * (\theta_1^i + \theta_2^i + \theta_{12}^0) - \mu_{w_s} - p_s^s \tag{19}$$

Taking Lemma 1 and (18) into account, one can see that if the farmer positively updates the mean of her beliefs regarding the profits of trait 1, the same will happen to the mean of the beliefs regarding the profits of trait 2. In addition (19) shows that the mean of the beliefs regarding the profits of the stacked variety will also be updated positively. In the same manner, if the mean of the beliefs were updated negatively for trait 1, the same will happen for trait 2 and for the stacked variety. Given the dependence of the stacked variety on both traits, the change in the mean of the beliefs for this variety, for a new observation, will always have a higher absolute value, as we can see in equation (20):

$$\frac{\partial \overline{\pi}_s^{B1}}{\partial n} = p^m * \left(\frac{\partial \theta_1^1}{\partial n} + \frac{\partial \theta_2^1}{\partial n}\right) \tag{20}$$

where *i=1* since equation (20) describes the mean of the posterior distribution after experience with trait 1. The second element of the expected utility contains the aggregate variance of profits, which captures the uncertainty regarding the value of the mean of the distribution of yields as well as the variance of yields and input costs that are caused by factors such as weather, pests and weed problems, i.e., risk. The aggregate variance for each type of variety is given by:

$$V_{ag}^i(\pi_c) = E(\pi_c - \overline{\pi}_c^{Bi})^2 = (p^m)^2 \sigma_c^2 + \sigma_{w_c}^2 \tag{21}$$

$$V_{ag}^i(\pi_1) = E(\pi_1 - \overline{\pi}_1^{Bi})^2 = (p^m)^2 (\sigma_1^2 + V_1^i) + \sigma_{w_1}^2 \tag{22}$$

$$V_{ag}^i(\pi_2) = E(\pi_2 - \overline{\pi}_2^{Bi})^2 = (p^m)^2 (\sigma_2^2 + V_2^i) + \sigma_{w_2}^2 \tag{23}$$

$$V_{ag}^i(\pi_s) = E(\pi_s - \overline{\pi}_s^{Bi})^2 = (p^m)^2 (\sigma_1^2 + V_1^i) + (p^m)^2 (\sigma_2^2 + V_2^i) + 2(p^m)^2 C_{12}^i + (p^m)^2 (\sigma_{12}^2 + V_{12}^0) + \sigma_{w_s}^2 \tag{24}$$

Equations (22)-(24) show that as experience increases and the variance of the beliefs go to zero, the aggregate variance converges to known risk, i.e., the variance of yields and input costs that is caused by factors such as weather, pests and weed problems.

It is worth noting that the uncertainty element is much higher for the stacked variety, given that it sums the uncertainty of both traits 1 and 2 as well as the uncertainty of μ_{12}. This greater uncertainty helps to explain why a farmer might adopt a single trait technology before adopting the stacked variety, because without experience with any of the traits in the technology, the uncertainty of the stacked variety is always higher than the uncertainty of a single trait variety. This points to one of the necessary conditions for a sequential adoption process, namely that with no information there are incentives to adopt a single part of the package rather than jumping directly into the package.

Proposition 1: *The uncertainty of the package technology will decrease as experience with an element of the package, n, increases.*

Proof: By Lemma's 2, and 3, and equation (9) in the following manner:

$$\frac{\partial V_{ag}^{1}(\pi_s)}{\partial n} = \frac{(p^m)^2 \partial V_1^1}{\partial n} + \frac{(p^m)^2 \partial V_2^1}{\partial n} + \frac{2(p^m)^2 \partial C_{12}^1}{\partial n} < 0 \qquad (25)$$

According to (9), Lemmas 2, and 3, each of the three terms that add up the change in the aggregate variance of the profits of the stacked variety in (25) is negative. Equation (25), then, shows that experience with an element of the package, such as a single trait variety, decreases the uncertainty associated with the elements of the package, which decreases the uncertainty of the entire package, or the stacked variety.

Proposition 2: *The change in variance is higher in absolute value for the stacked than for the single trait variety, i.e,* $\left|\frac{\partial V_{ag}^{1}(\pi_s)}{\partial n}\right| > \left|\frac{\partial V_{ag}^{1}(\pi_1)}{\partial n}\right|$ and $\left|\frac{\partial V_{ag}^{1}(\pi_s)}{\partial n}\right| > \left|\frac{\partial V_{ag}^{1}(\pi_2)}{\partial n}\right|$

Proposition 2 states that the uncertainty of the stacked variety will decrease more than the uncertainty of the single trait varieties. This proposition follows from equation (25), Lemma 2 and Lemma 3.

Proposition 3*: For an ascendant technology, the incentive to adopt a stacked variety is increasing in the number of trials with a single trait variety.*

Proof: The condition for the farmer to adopt a stacked variety is:

$$\begin{aligned} &\triangle_{sc} = EU_s - EU_c \geq 0 \quad \text{and} \\ &\triangle_{s1} = EU_s - EU_1 \geq 0 \quad \text{and} \\ &\triangle_{s2} = EU_s - EU_2 \geq 0 \end{aligned} \tag{26}$$

Looking at the impact of experience with trait 1 over the inequalities in (26), we have:

$$\frac{\partial \triangle_{sc}}{\partial n} = \frac{\partial \overline{\pi}_s^{B1}}{\partial n} - b\frac{\partial V_{ag}(\pi_s)}{\partial n} \tag{27}$$

$$\frac{\partial \triangle_{s1}}{\partial n} = p^m * (\frac{\partial \theta_2^1}{\partial n}) - b(\frac{(p^m)^2 \partial V_2^1}{\partial n} + \frac{2(p^m)^2 \partial C_{12}^1}{\partial n}) \tag{28}$$

$$\frac{\partial \triangle_{s2}}{\partial n} = p^m * (\frac{\partial \theta_1^1}{\partial n}) - b(\frac{(p^m)^2 \partial V_1^1}{\partial n} + \frac{2(p^m)^2 \partial C_{12}^1}{\partial n}) \tag{29}$$

The first term on the RHS of equations (27)-(29) reflects the impact of experience on the difference between the expected profits of the stacked variety and the expected profits of the alternative seed. The second term reflects the impact of experience on the difference between the aggregate variance of the stacked variety and the aggregate variance of the alternative seed. Taking into account Proposition 2, the second term will always be positive. That is, the aggregate variance of the stacked variety will always decrease more than the aggregate variance of the alternative seed. On the other hand, the sign of the first term will depend on whether the technology is ascendant or not.

With an ascendant technology, the beliefs regarding the mean yields of trait 1 are updated upwardly or do not change. As stated in Lemma 1 this implies that the mean of the beliefs regarding trait 2 are also updated upwardly or do not change. Thus, the first term of the RHS of equations (27)-(29) will be greater than or equal to zero.

Consequently, for an ascendant technology the impact of experience on the change in expected returns to the new technology ($\vartriangle_{sc}$, $\vartriangle_{s1}$ and $\vartriangle_{s2}$) will be greater than zero, increasing the incentive to adopt a stacked variety.

Proposition 4: For a non-ascendant technology, the incentive to adopt a stacked variety increases with the number of trials with a single trait variety if the reduction of the difference between expected yields is lower than the reduction of the difference between the variance of the expected utility of stacked versus non stacked seeds.

For a non-ascendant technology, where the beliefs are updated downwards, we have that the first term of the RHS of equations (27)-(29) is lower than zero while the second term of these equations is higher than zero. Hence, the impact of experience on $\vartriangle_{sc}$, $\vartriangle_{s1}$ and $\vartriangle_{s2}$ will be positive if the second term is higher in absolute value than the first term. This means that the incentive to adopt a stacked variety will increase with experience if the impact of experience on the difference between the variances of stacked versus non stacked seeds outweighs the negative impact of experience on the difference between the expected value of yields of stacked versus non stacked seeds.

Propositions 3 and 4 demonstrate the potential for a sequential adoption pattern by showing that, under certain conditions, experience with single trait varieties will increase the chances that a farmer adopts stacked varieties. Note that if (26) holds for n = 0; that is, if without a farmer observation on a single trait variety, the expected utility including both risk and uncertainty is higher for the stacked variety, then one would observe farmers jumping into the stacked variety directly rather than following a sequential adoption pattern through the single trait varieties.

Empirical Implementation

Econometrics of Sequential Adoption

Propositions 3 and 4 from the theoretical model show that, when the change in the expected value of yields is positive or if negative, higher than a certain threshold, the incentive to adopt a stacked variety will increase with the number of trials with the single trait variety. Formally, they present the following testable probability statement:

$$\Pr(j^* = s \mid n_i) > \Pr(j^* = s \mid n_j) \quad for \quad n_i > n_j \tag{30}$$ [8]

A standard estimation method of testing (30) would imply the use of a logit or probit model, but the standard version of these models imposes a linear relationship between the independent variable, *n*, and the latent variable, EU_j. Inspection of equations (8) and (9) shows that the latent variable, EU_j depends non-linearly on *n*, the level of experience, with higher levels of *n* having lower effects on expected utility than do lower levels of *n*.

Given this non-linearity in the role of experience, we estimate the adoption of a stacked variety using a semi-parametric specification.[9] It is as follows:

$$y_{it} = f(a_{it}) + \sum_T \alpha_T c_{tT} + v_i + \varepsilon_{it} \tag{31}$$

[8] Testing equation (30) provides evidence consistent with propositions 3 and 4, while its rejection would imply that a technology is non-ascendant and proposition 4 does not hold.
[9] An alternative specification, which can accommodate such non-linearity, would be to use survival analysis. Unfortunately, our data do not conform to the requirement of observing the characteristics of farmers at the beginning of the event, in this case the first introduction of GM seed traits in the mid 1990's, which is necessary for a robust survival analysis estimate.

where, y_{it} =1 if farm i adopts a stacked variety in year t, a_{it} is experience, measured as the years that have passed since the first year of adoption of a GM crop, and c_{tT} is a dummy variable for each chronological year, v_i is a fixed effect, and ε_{it} is the standard error term. The function $f(a_{it})$ has no predetermined parametric shape. The impact of experience occurs one year later: that is, if the farmer adopted GM corn for the first time in period t, a_{it} would be equal to zero and a_{it+1} would be equal to one. In addition to an individual's own experience, we expect that the expected utility of adopting a stacked variety may change over time. Including a time measure can capture rising levels of common information about the technology that a farmer access in addition to or as a substitute for their own experience.

A key challenge in estimating the impact of experience arises with its potential endogeneity. Unobservable variables, such as the farmer's ability, which could affect the adoption of stacked varieties, are also very likely to be correlated with early adoption. A possible solution implies the estimation of the likelihood of being an early adopter, which generates experience, as a function of the value of variables thought to cause early adoption such as land size and education, at the period at which the technology was first introduced. This, however, can cause an incidental parameters problem (Heckman). In order to estimate adoption as a function of experience while circumventing the incidental parameters problem, we use the Mundlak device in which we run an estimation that controls for the average value of experience of each household. In the current setting, this amounts to controlling for the first year of adoption.

Formally, since early adopters may possess characteristics that distinguish them from late adopters, it is possible for v_i to be correlated with a_{it}, which would introduce a bias in the estimation of $f(a_{it})$. Following Mundlak, we can define v_i as:

$$v_i = \sum_{j=0}^{6} \beta_j s_{ij} + z_i \tag{32}$$

where s_{ij} is equal to one if the sum of the variable a_{it} for farm *i* and over all the time periods is equal to *j*. While the fixed effects v_i are usually expressed as a function of the average of the independent variable, we use a more flexible expression, in which there is one dummy for each value of this average experience level. This is equivalent to using a dummy for the first year of adoption, with two exceptions. The first is due to the exclusion of a dummy that is equal to one if the farmer adopted a GM crop in 2000. This dummy is excluded because using a set of dummies that sum one for each observation will imply a collinearity problem in the context of this semi-parametric estimation. The second exception is related to the farmers who adopted for the first time in the last year of our data, 2006, or never adopted. These two groups of farmers are represented by the same dummy variable since they present the same value for the sum (and the average) of the variable a_{it}, namely zero. Using the Mundlak formulation to account for potential correlation of v_i and a_{it} the adoption of stacked varieties can now be expressed as:

$$y_{it} = f(a_{it}) + \sum_{T} \alpha_T c_{tT} + \sum_{j=0}^{6} \beta_j s_{ij} + z_i + \varepsilon_{it} \tag{33}$$

where having controlled for early adoption through s_{ij} we can now assume that $z_i + \varepsilon_{it}$ is not correlated with our variable of interest: a_{it}.

This semi-parametric specification offers the advantage of being highly flexible and of allowing us to know how the impact of experience changes as experience evolves. The estimation method follows Stock. Following Blundell and Duncan, we estimate confidence intervals for the non-parametric parameters using the bootstrap method. Additionally, the bootstrap replications allow us to correct for the bias, which characterizes nonparametric specifications.

Data:

The empirical analysis uses data gathered from a random sample of 738 corn farmers from Minnesota and Wisconsin from lists that the National Agricultural Statistics Service (NASS) maintains for their work on the agricultural census and other USDA data collection activities. These farmers were surveyed in 2006, and the questionnaire included retrospective questions on their use of GM corn varieties covering every year back to 2000. The empirical estimation exploits the variation in time of their adoption of GM varieties in order to assess the existence of a sequential pattern of adoption. We use data that covers the period 2000-2006 that contain information regarding the type of variety adopted each year. [10]

Descriptive Statistics:

Our first empirical illustration is to compare the timing of a farmer's first year of GM adoption with the use of stacked varieties in 2005, shown in table 1. The likelihood of

[10] Note that prior to 2005 we cannot distinguish the use of stacked varieties from the use of more than one single trait variety. This distinction is relatively unproblematic though, because in years when we have data on both stacked and multiple variety use, multiple variety use was very rare. For example, in 2005, the use of more than one variety was 2%, while the use of stacked varieties was 23%.

using stacked varieties in 2005 was twice as high (53% vs 27%) for farmers whose first year of GM adoption was in 2000-01 versus those adopting first in 2004-05. These descriptive results are consistent with a positive impact of experience through the reduction of uncertainty as well as with a technology that has proved to not be significantly worse than expected. Clearly, the results are not exclusive of other potential explanations. For one, this relationship might not imply causality because early adopters might have adopted the stacked varieties before accumulating any experience. Second, if there are variables correlated with early adoption, they might also shape farmer's use of stacked varieties. The econometric results in the next section exploit the panel data to account for the time the stacked variety was adopted and to isolate the impact of previous experience with single trait seeds.

Table 1
First year of Adoption and Type of Adoption in 2006 (Planned)

First year of adoption	% farms using stacked varieties 2006	% farms using Single trait varieties 2006	% farms not using GM seeds 2006
2000-01	53.4	40.4	6.2
2002-03	41.0	47.4	11.6
2004-05	29.6	50.0	20.3
2006	27.1	72.9	0.0
Total	41.4	48.2	10.4

Table 2 below shows the variables used in the semi-parametric estimation of equation (33) and the sample size of 4,157 observations. A farmer has adopted a stacked variety at 18% of the individual-year combinations in the data. On average across the dataset, 26% of the farmers adopted a GM variety in 2000 while 40% of the farmers adopted GM for the first time in 2006 or never adopted.

Semi-Parametric Estimation Results

Using the data summarized in table 2 we estimate the probability of adopting a stacked variety as expressed in equation (33). Since we do not impose any functional form for the impact of experience, it is possible to compare this impact for different groups.[11] In order to assess how this impact changes with time we have estimated equation (33) for the whole sample as well as for a sub-sample that includes observations for the later period only. To assess how the impact of experience changes with human capital, we have estimated equation (33) for two sub-samples: the first one contains farmers with no college education, while the second one contains farmers with college education[12].

Table 2
Descriptive Statistics

Variable	Description	Mean	Std. Dev.
y_{it}	Adopted more than one trait (yes = 1, No = 0)	0.18	0.38
a_{it}	Number of years since 1st year of adoption	1.22	1.81
t_t	Time Dummies for each chronological year 2002-2006 (each year represents ~ 17% of the sample)		
	Adopted a GM crop for 1st time in 2000 (Yes=1, No=0)	0.26	0.44
	Adopted a GM crop for 1st time in 2001 (Yes=1, No=0)	0.04	0.19
	Adopted a GM crop for 1st time in 2002 (Yes=1, No=0)	0.07	0.25
s_{ij}	Adopted a GM crop for 1st time in 2003 (Yes=1, No=0)	0.08	0.27
	Adopted a GM crop for 1st time in 2004 (Yes=1, No=0)	0.06	0.24
	Adopted a GM crop for 1st time in 2005 (Yes=1, No=0)	0.09	0.28
	Adopted a GM crop for 1st time in 2006 or never adopted (Yes=1, No=0)	0.40	0.49
	Number of Observations = 4,157		

[11] If, instead, we had assumed, for example, a linear relationship, a low coefficient on experience might have meant that farmers learn fast and later experience does not matter but it could have also meant that experience does not matter at any point. This vagueness would make it hard to compare the impact of experience for different groups. Even if we had modeled non linearities through the inclusion of a quadratic term, the groups comparison might have been misleading if $f(a_{it})$ were not a quadratic function. .

[12] When assessing how the impact changes with time we have not divided the sample in the early and late period because the range of the variable experience is too short in the early period.

Figure 1 shows the non-parametric component of the regression *f(a_{it})*, estimated according to equation (33), for the whole sample (2000-2006 period) as well as for the 2004-2006 period. The bands indicate the 90% confidence intervals around the predicted value lines[13]. The whole period regression includes as intercept the dummy for the year 2002, and the "later period" regression includes as intercept the dummy for the year 2005.

Figure 1 shows that, for the whole sample as well as for the later period, the likelihood of adopting more than one trait increases with experience. Experience is statistically significant for the whole period, since in this estimation there are some values of experience for which the confidence intervals do not intersect. These results are consistent with a positive impact of experience in reducing uncertainty and with Proposition 3 if a technology is ascendant or with the conditions stated in Proposition 4. On the other hand, the estimation for the later period presents a flatter slope and there are no values of experience for which the confidence intervals do not intersect, implying that experience is not a statistically significant predictor for the later period.

Table 3 shows the change in the likelihood of adopting more than one trait as experience increases by one year, for the whole sample as well as for the 2004-2006 period. For the whole sample, the impact of experience is lower at the outset, begins to increase at 2 years of experience and slows down at 5 years of experience. The impact of one additional year of experience on the likelihood of adoption ranges between 1.09% and 3.4%. This magnitude can be considered economically significant given that the

[13] As is common in the non-parametric literature, we chose to show the significance intervals at a 90% confidence level because the flexibility that characterizes non-parametric regressions comes at the cost of estimators with higher variances.

likelihood of adopting more than one trait for the individual-year combinations in the data is 18%, as shown in Table 2. As expected from Figure 1, Table 3 shows that the impact of experience is lower for the later period. Since the confidence interval of this difference lies in the positive range when the number of years of experience is equal to zero, three, four and five, we can assert that the difference in impacts between periods is statistically significant at the 95% level.

The flatter slope for own experience as well as the higher intercept in the later period likely reflect the higher degree of common knowledge about the technology, an issue we return to below. This first set of results supports the proposition that own experience shapes sequential adoption but diminishes in importance as own experience becomes less valuable as a way of acquiring information on the benefits of the technology.

Figure 1

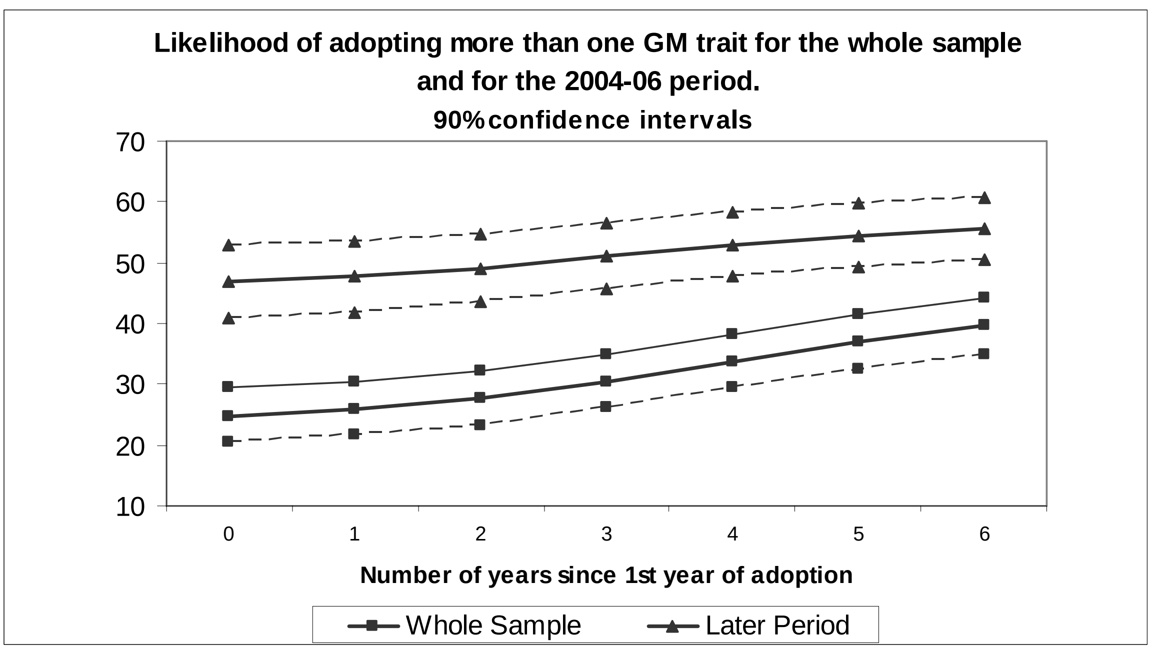


Table 3

Change in the likelihood of adopting more than one trait as experience increases by one year.

Experience *	Whole sample A	Later Period B	Difference A-B	95 % Confidence Interval of the Difference**	
0	1.09	0.75	0.34	0.02	0.63
1	1.80	1.32	0.48	-0.07	0.99
2	2.71	1.88	0.82	-0.03	1.55
3	3.37	1.97	1.40	0.40	2.18
4	3.27	1.55	1.73	0.78	2.46
5	2.53	1.06	1.47	0.64	2.09
Total	14.77	8.54	6.24	2.29	9.55

* Years since the first year of adoption
**Estimated using bootstrap

Next in figure 2, we consider the non-parametric component of the estimation for two sub-samples: one that contains farms whose operator has no college education and the other farms whose operator has a college degree or a higher educational level. In this figure, both functions include as intercept the dummy for the year 2002. For the sake of visual clarity, the confidence intervals are not included in this figure[14]. For both sub-samples, the 90% confidence interval that corresponds to six years of experience does not intersect with the confidence interval that corresponds to zero years of experience, implying that experience is statistically significant for both the college-educated and non college-educated farmers at the 90% level of confidence.

Table 4 shows the change in the likelihood of adopting more than one trait as experience increases by one year. The results show that the impact of experience for college-educated farmers is higher than the impact for non-college educated farmers at low levels of experience and lower at higher levels of experience. The difference between

[14] The confidence intervals are included in Aldana et al.

the impact of experience for these two sub-samples is negative and statistically different than zero when farmers have zero experience while being positive and statistically different from zero for farmers with four and five years of experience. Overall these results indicate that college- educated farmers tend to learn faster from own experience than less educated farmers.

Figure 2

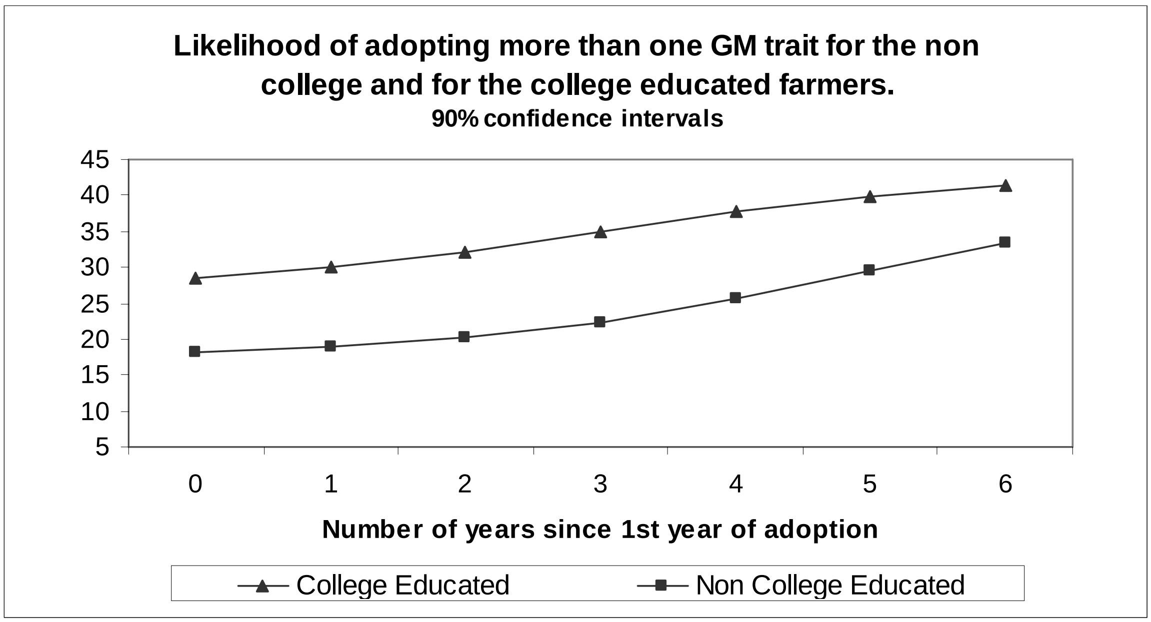


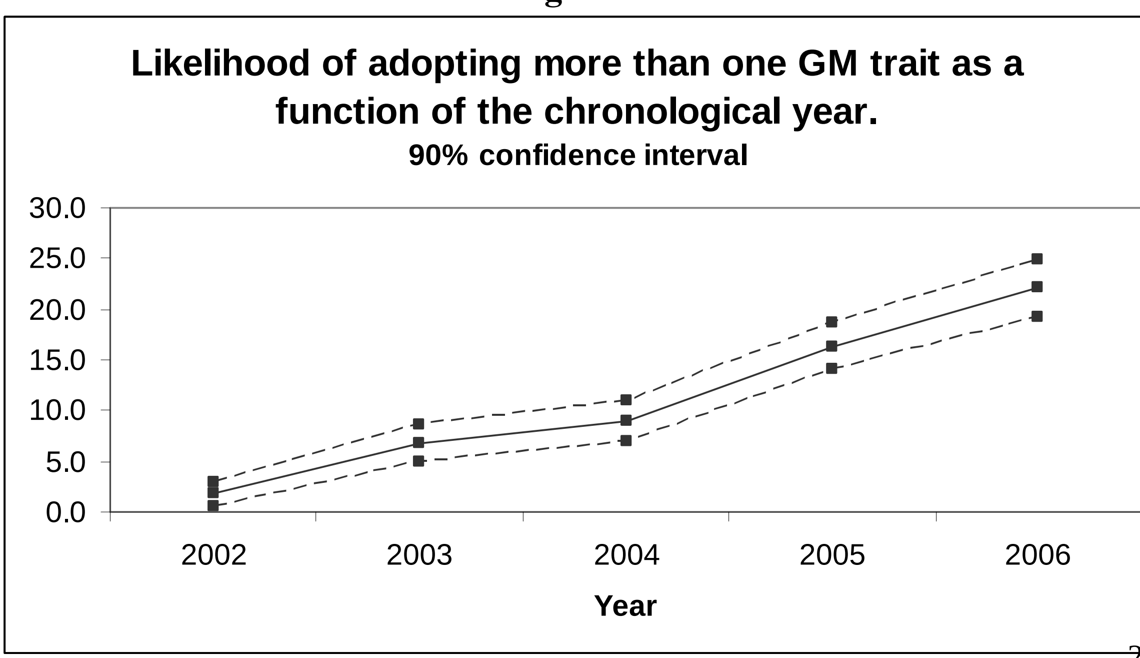
Table 4
Change in the likelihood of adopting more than one trait as experience increases by one year. Non College Educated vs College Educated Farmers.

Experience *	Non College Educated A	College Educated B	Difference A-B	95 % Confidence Interval of the Difference**	
0	0.70	1.46	-0.76	-1.20	-0.32
1	1.24	2.14	-0.90	-1.55	-0.18
2	2.16	2.74	-0.58	-1.66	0.59
3	3.31	2.80	0.51	-1.01	2.13
4	4.06	2.20	1.86	-0.04	3.66
5	3.89	1.38	2.51	0.35	4.09
Total	15.35	12.71	2.65	-3.88	9.58

* Years since the first year of adoption
**Estimated using bootstrap

In Figure 3, we use the time variables to investigate how aggregate levels of information affect the probabilities of individuals adopting the technology. It depicts the coefficients associated with the year dummies of the semi-parametric regression for the whole sample, i.e., α_t in specification (33). The coefficients on the year dummies have increased through time, with the diffusion of the technology, implying increases in the base levels of stacked variety adoption one would predict from a typical S-curve adoption pattern. Comparing the coefficient values in Figure 3 to Figure 1 provides insight into the relation between the impact of direct experience (Figure 1) and the impact of indirect experience (Figure 3). The results from Figure 1, of a positive impact of experience for the whole sample provide evidence of a sequential pattern adoption. However, the results shown in Figure 3 of a higher likelihood to adopt as time evolves, combined with the results on Table 3, which show that the impact of experience is lower for the later period, supports the notion that as time evolves and the uncertainty with a new technology dissipates, farms are more likely to jump directly to the use of stacked varieties and own experience loses some of its explanatory power.

Figure 3



In sum, the results from the semi-parametric estimation of the adoption of stacked crop varieties shows a sequential adoption process in which the initial experience with a single trait variety reduces the uncertainty and increases the probability of adopting a stacked variety. The estimates also show that farmers with high levels of education tend to learn faster from own experience and require fewer years of experience in order to adopt a stacked variety. In addition, the estimates show that with time, as the general level of knowledge in society about stacked varieties increases, this knowledge reduces uncertainty with the technology. Accordingly, over time the benefits of own experience and incentives for sequential adoption decline.

Conclusions

GM corn seed companies have innovated continuously with the introduction of new traits and, more recently, with the creation of stacked varieties, which combine more than one trait. In spite of its potential importance, no previous studies had examined the determinants of adoption of stacked varieties, particularly the learning dynamics and potential for sequential adoption in which farmers moved from use of GM varieties with a single trait to adoption of varieties with stacked traits. Our article fills this gap in the literature.

The Bayesian modeling framework is to our knowledge the first that models a sequential adoption process for a packaged good with correlated component technologies. It demonstrates the conditions under which experience with a single component of a package technology such as traits in corn seeds plays a role in reducing uncertainty regarding the profitability of the packaged (stacked) technology. As a consequence of this learning process, farmers' uncertainty associated with adopting a stacked variety will be lower conditional on having previously adopted one of the traits included in the seed.

The empirical results indicate that early adopters of GM will be more likely to adopt stacked varieties and benefit from the higher yields they offer. The empirical evidence presented above shows that the likelihood of adopting a stacked variety increases with experience, measured as the number of years that have passed since the first year of adoption of a GM technology. The results also show that the impact of experience decreases with time as the technology diffuses, and that farmers are more likely to jump directly to the use of stacked varieties in the later years. They also show that more educated farmers tend to learn faster from own experience.

The sequencing of adoption of a package technology has important implications for both the introduction of new technologies to the market and the industrial organization of the market for new technologies. In terms of the introduction of new technologies, these results suggest strong incentives for technology sellers or promoters to help early adoption of component parts of a stacked technology as a strategy for eventual adoption of the package technology. In contrast, efforts to push adoption of stacked technologies that ignore the uncertainty that drives sequential adoption processes may run into difficulty. The sequential nature of adoption may drive differential pricing strategies among technology sellers and this could have profound implications for competition and the ability of small producers to compete. These industrial organization questions are left for future research.

References

Aldana, U., J. Foltz, B. Barham and P. Useche. 2010. "Sequential Adoption of Package Technologies The Dynamics of Stacked Trait Corn Adoption." *Agricultural and Applied Economics. Staff Paper Series*. UW-Madison.

Alexander, C. 2006. "Farmer Decisions to Adopt Genetically Modified Crops." *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources* 1: 1-9.

Anderson, J. 1977. *Agricultural Decision Analysis*, Iowa City: Iowa University Press.

Blundell, R. and A. Duncan. 1998. "Kernel Regression in Empirical Microeconomics." *The Journal of Human Resources*, 33: 62-87.

Byerlee, D. and E. Hesse de Polanco. 1986. "Farmers' Stepwise Adoption of Technological Packages: Evidence from the Mexican Altiplano." *American Journal of Agricultural Economics*, 68: 519-527.

Conley, T.G. and C.R. Udry. 2004. "Learning About a New Technology: Pineapple in Ghana." Economic Growth Center, Yale University. Working Paper # 817.

Feder, G., R. Just, and D. Zilberman. 1985. "Adoption of Agricultural Innovations in Developing Countries: A Survey." *Economic Development and Cultural Change* 33 : 255-298.

Feder, G. and R. Slade. 1984. "The Acquisition of Information and the Adoption of Technology." *American Journal of Agricultural Economics,* 66: 312-320.

Feder, G. and G. O'Mara. 1981. "Farm Size and the Diffusion of Green Revolution Technology." *Economic Development and Cultural Change*, 30: 59-76.

Fernandez Cornejo, J., S. Daberkow and W. McBride. 2001. "Decomposing the Size Effect on the Adoption of Innovations: Agrobiotechnology and Precision Agriculture." *Agbioforum,* 4: 124-136.

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

Griliches, Z. 1957. "Hybrid Corn: An Exploration in the Economics of Technological Change." *Econometrica* 25:501-22.

Heckman, J.J. 1981. "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process." In:

C.F. Manski and D. McFadden eds. 1981. *Structural Analysis of Discrete Data with Econometric Applications,* Cambridge, MA: MIT Press.

Hubbell, B.J. M.C. Marra and G.A. Carlson. 2000. "Estimating the Demand for a New Technology: Bt Cotton and Insecticide Policies." *American Journal of Agricultural Economics*, 82:118–32.

Leathers, H and M. Smale. 1991. "A Bayesian Approach to Explaining Sequential Adoption of Components of a Technological Package." *American Journal of Agricultural Economics*, 73: 734-742.

Lindner, R., A. J. Fischer and P. Pardey. 1979. "The Time to Adoption." *Economic Letters* 2: 187-90.

Magnac, T. 2000. "Subsidized Training and Youth Employment: Distinguishing Unobserved Heterogeneity from State Dependence in Labour Market Histories." *The Economic Journal,* 110: 805-837.

Mundlak, Y. 1978. "On the Pooling of Time Series and Cross Section Data." *Econometrica* 46: 69-85.

Rogers, E. 1986. *Diffusion of Innovation,* New York: The Free Press.

Stock, J. 1989. "Nonparametric Policy Analysis." *Journal of the American Statistical Association*, 84: 62-87.

Stoneman, P. 1981. "Intra-Firm Diffusion, Bayesian Learning and Profitability" *The Economic Journal*, 91: 375-388.

Useche, P., B. Barham, and J. Foltz. 2009. "Integrating Technology Traits and Producer Heterogeneity. A Mixed Multinomial Model of GM Adoption." *American Journal of Agricultural Economics*, 91: 444-461.