A Dynamic Adoption Model with Bayesian Learning: An Application to U.S. Soybean Farmers

Xingliang Ma[§]

and

Guanming Shi

Abstract:

Adoption of agricultural technology is often sequential, with farmers first adopting a new technology on part of their lands and then adjusting their use of the new technology in later years based on what was learned from the initial partial adoption. Our paper explains this experimental behavior using a dynamic adoption model with Bayesian learning in which forward-looking farmers take account of future impacts of their learning from both their own and their neighbors' experiences with the new technology. We apply the analysis to a panel of U.S. soybean farmers surveyed from 2000 to 2004 to examine their adoption of the genetically modified (GM) seed technology. We compare the results of the forward-looking model to that of a myopic model, in which farmers maximize current benefits only. Results suggest that farmers in our sample are more likely to be forward-looking decision makers. The myopic model underestimates the value of early adoption for forward-looking farmers, and predicts lower adoption rates at the beginning of our study period. We also find evidence that farmers tend to rely more on learning from their own experience than on learning from their neighbors.

Keywords: Technology adoption, Bayesian learning, structural estimation

JEL codes: D83, Q16, O31, O33

[§] Corresponding author. Email: xingliang.ma@cgiar.org. Postdoctoral fellow at the International Food Policy Research Institute and associate professor at the Department of Agricultural and Applied Economics, University of Wisconsin, Madison. This research was funded in part by USDA-NRI grant #144-QS50 and by a USDA CSREES Hatch grant #WIS01345.

1 Introduction

Many researchers have modeled agricultural technology adoption as a binary choice problem: with farmers choosing either to adopt a new technology or not to adopt the technology at all (e.g. Cameron, 1999; Barham et al., 2004; Useche, Barham and Foltz, 2009). This binary choice assumption allows researchers to use the Logit or multinomial Logit method to analyze the adoption process. However, in many cases farmers may more likely try new technologies sequentially or "stepwise." Farmers may choose to apply a new technology to part of their lands first, and then adjust adoption practices in later years after observing outcomes from the earlier partial adoption. For example, during the "Green Revolution" era, farmers often initially experimented with the new seed varieties, fertilizer, and other new agricultural practices on offer, adopting them only partially at first. Cummings (1975) observed that "Farmers … experiment with recommendations, often adopting them in stages rather than as a complete package." (Cummings, 1975, p. 24). Foster and Rosenzweig (1995) and Munshi (2004) both acknowledge the experimental behavior of India farmers on optimal input use during their adoption of high yield varieties (HYVs) in 1968-1970. A similar pattern is observed in adoption by U.S. farmers of genetically modified (GM) seeds since the mid-1990s. Farmers rarely switch all their land from conventionally bred seeds to GM seeds immediately. Rather, the adoption process is gradual, and farmers who adopt GM seeds may only achieve full or partial adoption even after 15 years since the inception of the GM seed technology.[1]

What factors drive such an adoption pattern? Technology adoption is a complex and dynamic process, involving risk management, learning and investment adjustment (Griliches,

[1] Note the difference between sequential adoption and partial adoption in equilibrium. Partial adoption in equilibrium is due to farmland heterogeneity. Part of their land may not be suitable for the new technology, thus it may be optimal to adopt partially. Sequential adoption refers to the process: It takes several years to reach the "equilibrium" level of adoption, either full or partial.

1957; Barham et al., 2004; Aldana et al., 2011). Facing a new technology and with limited information, a risk-averse farmer may perceive a high risk with the technology and, in general, will not adopt it to all her land immediately. The gradual adoption pattern results from the gradual flow of information and gradual change of farmers' perceptions about a new technology. Furthermore, such a gradual adoption process is complicated if we also consider farmers' forward-looking behavior.

A farmer may adopt a new technology to part of her land even if this adoption is not optimal for the current time period believing that experience garnered from current adoption will give her valuable information on the new technology to assist her in making better future decisions. Forward-looking farmers take future impacts, both negative and positive, into account when making current adoption decisions.

Such impacts may be interpreted as adoption externalities. For example, if a new technology entails uncertainty and potential risks in crop yield or farm profitability, partial adoption carries a positive externality since, as farmers experiment with partial adoption today, they gain improved knowledge about the new technology's profit distribution through learning. By internalizing such positive externalities, farmers may find it optimal to partially adopt a new technology, although the profit with partial adoption may be less than that with no adoption for the current period.

If, however, a new technology induces deteriorating soil quality, or resistance build-up in weeds or insects, potentially increasing risk and uncertainty of future crop production, then current adoption may have future negative externalities. Forward-looking farmers may find partial adoption more desirable than full adoption, although in this case current profits may be higher with full adoption than with partial adoption.

In this paper, we define positive adoption externality as “adoption benefits”, and negative externality as “adoption costs”. If farmers fail to account for adoption benefits/costs in decision making, dubbed as “myopic”, then they may make sub-optimal adoption decisions, either under-adopting in the case of positive adoption externalities, or over-adopting in the case of negative adoption externalities. In most cases, farmers are assumed to be forward-looking as long as they stay in the farming business. Myopic farmers are rare, but they might be found in those with an expiring land tenure contract.

Much research has been conducted to understand technology adoption in agriculture. Following Griliches (1957), early agricultural technology adoption literature focuses on how farmer characteristics and farmland heterogeneity affect adoption decisions under a static setup. For example, Feder, Just and Zilberman (1985) survey the literature on agricultural technology adoption and suggest that farm size, risk and uncertainty, human capital, labor availability and credit constraints contribute to differences in adoption. Useche, Barham and Foltz (2009) employ a mixed multinomial Logit model to investigate the effect of heterogeneity in both farmers and GM corn seeds on farmers’ adoption decisions of GM technology. Their results show that farmers adopt different types of GM seeds according to their preferences for different traits embedded in the seeds.

Recent literature recognizes the dynamic nature of the adoption process and incorporates the learning component into adoption models (e.g., Besley and Case, 1994; Foster and Rosenzweig, 1995; Baerenklau, 2005). Both Besley and Case (1994) and Foster and Rosenzweig (1995) model farmers’ adoption of high-yielding seed varieties with learning in India during the Green Revolution. Comparing models with various assumptions on learning behavior of farmers, Besley and Case (1994) find that the myopic model performs least well, and the cooperative

learning model, in which farmers learn collectively within a village, performs the best in predicting the technology diffusion path. Foster and Rosenzweig (1995) explicitly model farmers' learning of optimal input usage and compare the effect of self-learning versus learning from neighbors. Both papers confirm that imperfect knowledge of the new technology inhibits adoption and farmers' learning can significantly reduce uncertainty. Baerenklau (2005) builds a similar adoption model with a focus on risk preferences, learning and peer-group influences. He applies the model to a group of Wisconsin dairy farmers and finds that risk preferences and learning are key factors driving technology adoption, and that peer-group influence plays a less important role than self-learning.

In this paper, we construct a continuous choice dynamic model in which forward-looking farmers learn about a new technology by applying it to part of their land. Based on their beliefs regarding the new technology's risk, farmers solve a finite period dynamic programming problem to choose the amount of land to allocate to the new technology in each time period. Unlike previous literature that focuses on the learning of mean profit (e.g., Foster and Rosenzweig, 1995; Besley and Case, 1994), our model focuses on farmers' perceived profit variance associated with the new technology. Moreover, our structural model estimation recovers all the model parameters by searching within the whole parameter space, which differs from the previous dynamic adoption literature that either relies on reduced form estimation recovering only part of the parameters (e.g., Foster and Rosenzweig, 1995), or conducts the parameter search within a limited parameter space (e.g. Besley and Case, 1994; Baerenklau, 2005).

Our model is applied to a panel of U.S. soybean farmers from years 2000 to 2004. The two major types of seed technologies in the U.S. soybean seed market are conventionally bred seeds and GM herbicide tolerant seeds that allow farmers to apply specific herbicides without

crop damage. We estimate both myopic and forward-looking models, and compare predicted adoption patterns from both models with actual adoption behavior. We find that the mean squared errors (MSE) in the forward-looking model are smaller than in the myopic model, suggesting that farmers in our sample may behave in accordance with the forward-looking assumption. Applying the myopic model to the forward looking farmers generates lower predicted adoption rates during the early years, suggesting that early adoptions generate future adoption benefits, which may be captured through learning. Results also show that farmers in our sample learn more from their own experience than from their neighbors, consistent with current literature on social learning (e.g. Besley and Case, 1994; Foster and Rosenzweig, 1995; Baerenklau, 2005). However, we also find that when applying the "wrong" myopic model to forward-looking farmers, the model generates upward biased results for the noise in learning from neighbors. Thus, research using myopic models to examine the role of social learning may suffer from model misspecification, and therefore underestimate the value of social learning.

The paper is organized as follows. Section 2 presents the model, in which we specify the distribution of returns from two technologies: a conventional technology and a new technology, and construct farmers' Bayesian learning process accordingly. We describe the data in Section 3. In Section 4, we explain the estimation strategies for both the myopic model and the dynamic forward-looking model. Sections 5 and 6 present the estimation results and conclusion.

2. An Adoption Model with Bayesian Learning

Suppose farmers are faced with two technologies: an existing conventional technology (old) and a newly developed technology (new). Assume that profits of both technologies are random, i.e., both technologies are risky assets for farmers. If farmers are myopic, they will choose the adoption rate to maximize current net benefits only. However, if farmers are forward-looking,

they will make a sequence of adoption decisions to maximize total discounted net benefits over time.

2.1 A Mean-Variance Framework

Suppose the total profit π for each farmer is normally distributed, then expected utility $u(\pi)$ can be expressed as a function of the mean and the variance of the profit (Huang and Litzenberger, 1988, p.61). So, for farmer i

$$u_i(\pi) = f(E[\pi], \sigma^2(\pi)),$$

where $E[\pi]$ and $\sigma^2(\pi)$ are the mean and the variance of the total profit, respectively. Assume $f(\cdot)$ is a linear function,

$$f(E[\pi], \sigma^2(\pi)) = E[\pi] - \frac{1}{2}\rho_i \sigma^2(\pi),$$

where ρ_i is a measure of farmer i 's degree of risk aversion and is specified as

$$\rho_i = \beta_0 + \frac{\beta_1}{A_i},$$

where A_i is the farm size of farmer i, and β_0, β_1 are corresponding parameters[2]. Then, write the expected utility of farmer i as:

$$u_i(\pi) = E[\pi] - \frac{1}{2}\rho_i \sigma^2(\pi) = E[\pi] - \frac{1}{2}\left(\beta_0 + \frac{\beta_1}{A_i}\right)\sigma^2(\pi). \tag{1}$$

Next we specify the distributions of profits from the old and the new technology.

2.2 Distribution of Returns

Assume the technologies are seeds: the old technology is conventionally bred seed and the new technology is GM seed. The profit per unit of land is assumed to be normally distributed. For

[2] In this specification, we allow a farmer's risk attitude to be affected by farm size, an indicator of her wealth status. We test whether such a correlation is significant or not in the empirical study.

conventional seed, since it has been planted for many years, assume its distribution is known to any farmer i at time t as:

$$\pi_{ict} \sim N(\mu_{ict}, \sigma_{ict}^2).$$

where μ_{ict} is the average profit of conventional seed for farmer i at time t, and σ_{ict}^2 is its variance. For GM seed, the profit at time t for farmer i is assumed to be

$$\pi_{igt} = \mu_{ig} + \varepsilon_{igt},$$

where μ_{ig} is the average profit of GM seed for farmer i, and ε_{igt} is a random term following an independent and identical normal distribution with mean zero and time invariant variance σ_{ε}^2, $\varepsilon_{igt} \sim N(0, \sigma_{\varepsilon}^2)$. The random term ε_{igt} may include the impact of unpredictable weather such as rain fall, and unobserved factors such as soil conditions and individual farmer characteristics. Farmers, because of imperfect knowledge of the new technology, only perceive the average profit of GM seed μ_{ig} with uncertainty, and their beliefs follow a normal distribution $\mu_{ig} \sim N(\zeta_{igt}, \varphi_{igt}^2)$, which can be updated over time based on their own experience and information obtained from their neighbors.

Specifically, farmers' learning process on the GM average profit μ_{ig} is as follows: at time zero, farmer i receives exogenous information on the GM average profit as ζ_{ig0}, for which farmer i believes its accuracy can be measured as φ_{ig0}^2. This information may come from agronomists and agricultural extensionists, or from farmers' own observations of pest/weed infestation in past years and possible effectiveness of new GM traits. At time1, if it is profitable based on this prior information, farmer i may experiment with GM seed on part of her land, and then update her beliefs on both parameters to ζ_{ig1} and φ_{ig1}^2, using the information learned from

the field experiment. Meanwhile, farmer i observes her neighbors' behavior and may also incorporate this information to update ζ_{ig1} and φ_{ig1}^2. At time 2, similarly, farmer i decides whether or not to experiment with the new seed, how many acres she should allocate to GM seed, and update her beliefs to get a new set of ζ_{ig2} and φ_{ig2}^2 accordingly. This learning process keeps going, and in each time period farmers acquire additional information about GM seed and, therefore, become more certain of the profitability of GM seed.

Both forward-looking farmers and myopic farmers learn the GM average profit μ_{ig} in the same way as explained above; however, forward-looking farmers consider future impact in their current decision making, while myopic farmers do not. Since they have different attitudes towards the future, they may value this learning experience differently. If learning has a positive externality in the future, forward-looking farmers will do more experiments than myopic farmers in the early time period to capture future benefits from early learning experiences. Similarly, if learning has a negative externality, forward-looking farmers will internalize that information and have less learning experiments than myopic farmers in the early stage of adoption.

In the following, we specify the detailed learning mechanism of the GM average profit and the distribution of the profit for conventional seed.

Update of the Perceived GM Profit Variance

With the uncertainty of the GM average profit, the total variance of the profit from planting GM seed for farmer i at time t is

$$\sigma^2(\pi_{igt}) = \sigma^2(\mu_{ig}) + \sigma^2(\varepsilon_{igt}) = \varphi_{igt}^2 + \sigma_\varepsilon^2, \tag{2}$$

where the profit variance induced by the disturbance ε_{igt}, σ_ε^2, is assumed to be a known parameter for farmers; the first term φ_{igt}^2, farmer i's perceived variance of the GM average profit

or the uncertainty associated with adopting GM seeds, is unknown for farmers but can be reduced by learning either from their own experiments or from the observation of their neighbors' experience. Since GM is a new technology, the uncertainty is high and farmers may initially perceive a high variance with its profitability. This perceived variance may decrease over time if farmers learn about this new technology by experimenting on part of their land and/or by communicating with their neighbors. Figure 1 illustrates a possible path of the perceived variance of GM profit over time with a constant belief of the mean: at time 0, farmers' perceived variance of μ_{ig} is high; With experiments over time, farmers become less and less uncertain about μ_{ig} and their perceived variance φ_{igt}^2 become lower as time t increases.

If the learning process of each farmer follows a Bayesian setup, then farmer i updates her perceived profit variance of GM seed in the following way[3]

$$\varphi_{igt+1}^2 = \frac{1}{\frac{1}{\varphi_{igt}^2} + \frac{G_{it}}{\sigma_\varepsilon^2} + \frac{G_{-it}}{\sigma_\varepsilon^2 + \sigma_\xi^2}}, \tag{3}$$

where G_{it} is farmer i's total adopted units of land of the GM seed at time t, G_{-it} is the average adopted total units of land of her neighbors, σ_ξ^2 is the additional variance or noise in farmer i's learning from neighbors.

[3] See Appendix A for the detailed derivation.

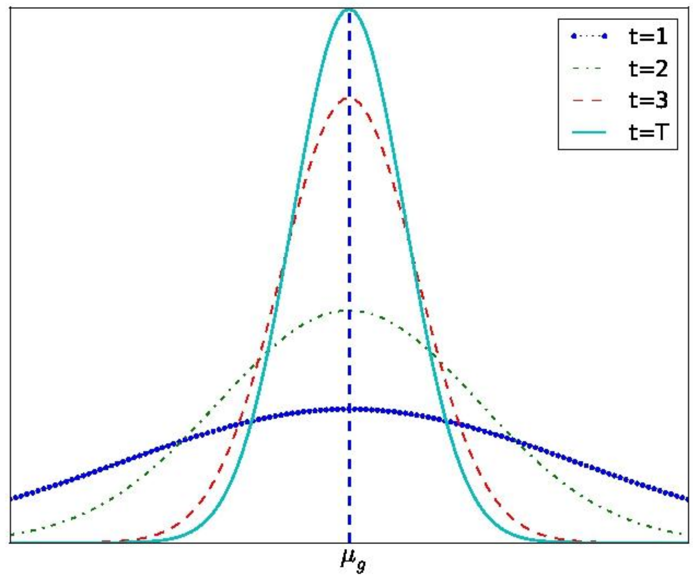

Figure 1: Update of the Perceived GM Profit Variance

This formula implies that if farmer i does not adopt any GM seed at time t, and does not obtain any information from her neighbors, her belief toward the variance of the GM profit stays the same as it was at time $t-1$. If farmer i experiments with GM seed on part of her land at time t, then the more she plants the GM seeds (increase in G_{it}), the more she will learn about μ_g (decrease in σ^2_{igt+1}). And if this farmer lives in a region with high adoption rates among her neighbors (increase in G_{-it}), she will also have a better knowledge of the GM technology (decrease in σ^2_{igt+1}).

However, the information farmer i could get from neighbors may carry additional noise as compared to information obtained from her own experience ($|\frac{\partial \varphi^2_{igt+1}}{\partial G_{it}}|>|\frac{\partial \varphi^2_{igt+1}}{\partial G_{-it}}|$). The noise in neighborhood information may come from two sources: 1) some information may get lost during the communication; and 2) if the average GM profit depends on farmers' individual

characteristics, as argued by Manski (1993) and Munshi (2004), the information from the neighbors may be biased and not applicable to her own case.

Variance of the Profit for Conventional Seed

The variance of profit from planting conventional seed is known for farmers. However, since conventional seed may be vulnerable to uncertain events such as pest infestations or weeds, we assume its variance depends on a random state variable z_t which follows an AR(1) process. For farmer i at time t, the variance of the profit from planting conventional seed, σ_{ict}^2, is

$$\sigma_{ict}^2 = \sigma_{ict}^2(z_t), \text{ and}$$
$$z_t = \lambda z_{t-1} + \nu_t, \quad \nu_t \sim N(0, \sigma_\nu^2)$$

where ν_t is the white noise added to the AR(1) process in each time period. Assume σ_{ict}^2 is a linear function of z_t

$$\sigma_{ict}^2 = \sigma_{ict}^2(z_{it}) = \gamma_0 + \gamma_1 z_t = \gamma_0 + \gamma_1(\lambda z_{t-1} + \nu_t). \tag{4}$$

Assume γ_1 is positive, meaning that a higher variance of the random state variable brings a higher variance of the profit from the conventional seed, and, therefore, a lower expected utility to a risk-averse farmer.

Mean Profit

In reality, farmers update their beliefs on both the mean and the variance of the average GM profit μ_{ig}, using the information of the realized profits from either their own adoption or from their neighbors' adoption.[4] However, we do not have the information on the actual profits in our empirical study, so we assume that farmers' beliefs on the GM mean profit ζ_{igt} is constant, i.e., initially farmers receive an unbiased estimator of the mean on perceived profit of GM seed, and

[4] See Appendix A for the derivation of the updating rules for both the mean and the variance.

then update their beliefs on the accuracy (the variance φ_{igt}^2) in the later time periods.[5]

Moreover, we assume that there is heterogeneity in farm land characteristics and that a farmer can conceptually arrange all her lands in such a way that the suitability of the land for planting GM seeds is decreasing. This suitability for GM seeds may be related to soil conditions, land quality, infestation vulnerability, or other factors. Suppose farmer i owns a total of A_i units of land plots and for each plot the mean profit of conventional seed is η_c, the difference between the unbiased belief of GM mean profit and the conventional mean profit for the k^{th} plot, $\Delta\mu_i^k$, is

$$\Delta\mu_i^k = \zeta_{ig}^k - \eta_c = \eta_{ig}(X_i) - \eta_{gc}\frac{k}{A_i} \qquad \text{where } k = 1, 2, \ldots, A_i.$$

where ζ_{ig}^k is farmer i' s belief of the GM profit from the k^{th} plot, η_{ig} is the upper bound of the profit difference, which is a linear function of farmer i's characteristics X_i, i.e., $\eta_{ig} = \eta_g + cX_i$. Assume $\eta_{gc} > 0$, i.e., the mean profit difference between the GM and conventional seeds is decreasing in k.

If farmers' adoption decisions are made based on comparing mean profits only, without forward looking, the optimal adoption rate is then determined by the intercept η_{ig} and the slope η_{gc}. Figure 2 plots scenarios where the optimal adoption rate α_{it} (defined as $\alpha_{it} \equiv \frac{G_{it}}{A_i}$) can be zero (line C: no adoption), one (line A: full adoption) or between zero and one (line B: partial adoption at 40%).

[5] We impose this restriction to facilitate the empirical analysis for the U.S. soybean market. Our assumption may not be overly restrictive. As agronomists point out (Hurley, Mitchell and Rice, 2004), in general GM technology does not increase but insures potential yield, which implies that the benefit of adopting GM seed may be mainly the reduced profit variance.

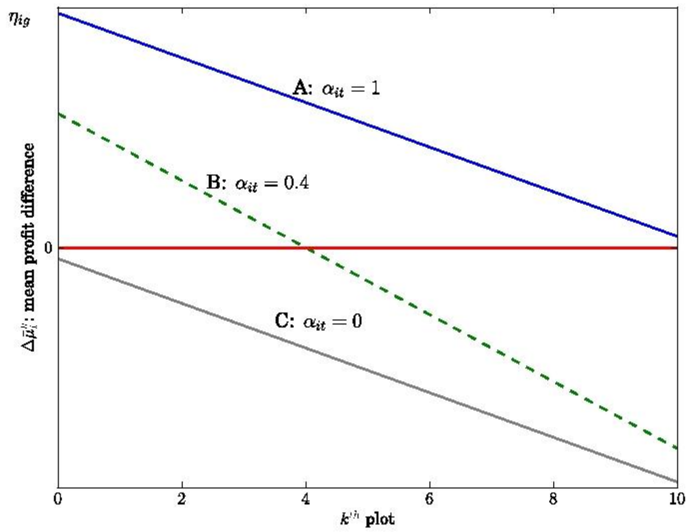

Figure 2: Mean Profit of GM Seed

Therefore, suppose farmer i adopts a total of G_{it} plots of GM seed at time t, the total mean profit she could get is[6]

$$E[\pi_{it}] = \sum_{k=1}^{G_{it}} \left(\eta_c + \eta_{ig} - \eta_{gc} \frac{k}{A_{it}} \right) + \sum_{k=G_{it}+1}^{A_{it}} (\eta_c) = \left(\eta_c + \eta_{ig}\alpha_{it} - \frac{1}{2}\eta_{gc}\alpha_{it}^2 \right) A_{it}.$$

2.3 Adoption Process

Assume independence of profits from different land plots. Based on previous specifications, the mean and variance of the total profit for farmer i at time t is

$$E[\pi_{it}] = \left(\eta_c + \eta_{ig}\alpha_{it} - \frac{1}{2}\eta_{gc}\alpha_{it}^2 \right) A_{it} \tag{5}$$

$$\sigma^2(\pi_{it}) = A_{it}^2 \left(\alpha_{it}^2 (\varphi_{igt}^2 + \sigma_\varepsilon^2) + (1-\alpha_{it})^2 \sigma_{ict}^2 \right), \tag{6}$$

where φ_{igt}^2 and σ_{ict}^2 are specified in Equations (3) and (4).

The current payoff at time t for farmer i is

[6] See Appendix B for a detailed derivation of both the mean and variance of the total profit.

$$
\begin{aligned}
u_{it} &= E[\pi_{it}] - \frac{1}{2}\rho_i \sigma^2(\pi_{it}) \\
&= \left(\eta_c + \eta_{ig}\alpha_{it} - \frac{1}{2}\eta_{gc}\alpha_{it}^2\right) A_{it} - \frac{1}{2}\rho_i A_{it}^2 \left(\alpha_{it}^2(\varphi_{igt}^2 + \sigma_\varepsilon^2) + (1-\alpha_{it})^2 \sigma_{ict}^2\right) \qquad (7) \\
&= u_{it}\left(\alpha_{it}^2, \varphi_{igt}^2\left(G_{it-1}, G_{-it-1} \mid \sigma_\varepsilon^2, \sigma_\xi^2\right), \sigma_{ict}^2\left(z_t, \nu_t \mid \gamma_0, \gamma_1\right), A_{it} \mid \eta_c, \eta_{ig}, \eta_{gc}, \rho_i\right) \\
&\equiv u_{it}\left(\alpha_{it}, S_{it} \mid \Theta\right),
\end{aligned}
$$

where S_{it} is the state variable, which includes the current belief of the GM profit variance φ_{igt}^2, the profit variance of conventional seeds σ_{ict}^2, and the total soybean acreage A_{it}. The set Θ is the parameter space of the model, defined as $\Theta \equiv \{\sigma_\varepsilon^2, \sigma_\xi^2, \sigma_\nu^2, \gamma_0, \gamma_1, \eta_c, \eta_{gc}, \eta_{ig}, \rho_i\}$.

It is commonly observed that technology diffusion follows an S-curve, i.e., the new technology spreads at an increasing rate during the early period, then its adoption rate growth slows down gradually, and eventually the adoption rate remains at a constant level. This adoption pattern also holds for GM soybean seed in the US. After its introduction in the mid-1990s, GM soybean seed spread across the US rapidly. In about 10 years, especially after 2004, the GM adoption rate became flat. In light of this, we model the dynamic adoption problem as a finite period dynamic model, i.e., farmer i chooses a sequence of actions $\{\alpha_{il}\}_{l=t,t+1,\ldots,T}$ to maximize total discounted expected utility from time t to the steady state time period T,

$$
V_{it} = \max_{\{\alpha_{il}\}_{l=t}^{T}} E_t \sum_{l=t}^{T} \delta^{T-l} u_{il}(\alpha_{il}). \qquad (8)
$$

The steady state could be a complete switch to GM seed, or a partial adoption that farmers wish to maintain. With current payoff defined as in Equation (7), the Bellman Equation is

$$
V_{it}(S_{it}) = \max_{\alpha_{it}} \{u_{it}\left(\alpha_{it}, S_{it} \mid \Theta\right) + \delta E V_{it+1}(S_{it+1} \mid S_{it})\}. \qquad (9)
$$

For forward-looking farmers, the optimal adoption rates in each time period are solved by backward induction based on this Bellman equation.

3. Data

To illustrate our method empirically, we apply the model to the adoption of GM soybean in the U.S. market. The soybean market is chosen for two reasons. First, as in the theoretical model, it comprises only two technologies, conventionally bred seeds and herbicide tolerant GM seed. Second, adoption of GM soybean seed in the U.S. reached an average adoption rate at 85% in year 2004, and adoption growth has slowed down since then, which justifies the finite period assumption in our model.

The empirical analysis is based on extensive survey data collected by dmrkynetec (hereafter DMR). The DMR data, obtained from a stratified sample of U.S. soybean farmers surveyed annually, provide farm-level information on seed purchases, acreage, seed types, and seed prices. We identify a panel of 432 farmers surveyed from 2000 to 2004 out of a total of 11,060 farmers in the DMR data. Note that observations from 2000 are used as a benchmark only. Figure 3 shows the average adoption rate of GM soybean seeds of these 432 farmers from 2000 to 2004, and the average adoption rate of GM soybean seeds from the whole DMR data and from USDA NASS data during this time period.[7] The sample average adoption rate follows a similar pattern as the DMR population and the USDA NASS population. It suggests that farmers in our sample do not differ from farmers in the population as a whole in terms of adoption behavior.

To avoid the complication caused by farmers' switching between soybean and other crops across the years, we focus on farmers with relatively constant soybean acreage over this time period.[8] After screening, 348 farmers are included in our sample for analysis. Figure 4 shows that most of these 348 farmers are located in the Midwest of U.S.

---

[7] USDA Data are collected from the official website at http://www.nass.usda.gov/, in the reports on "acreage" from year 2000 to 2004.

[8] We construct a farm size variation measure by dividing the standard deviation of the farm size by its mean. We dropped those farmers with greater than 30% variation. We also tried screening at other levels including 10%, 20%, and 40%, and the final estimation results are qualitatively similar. .

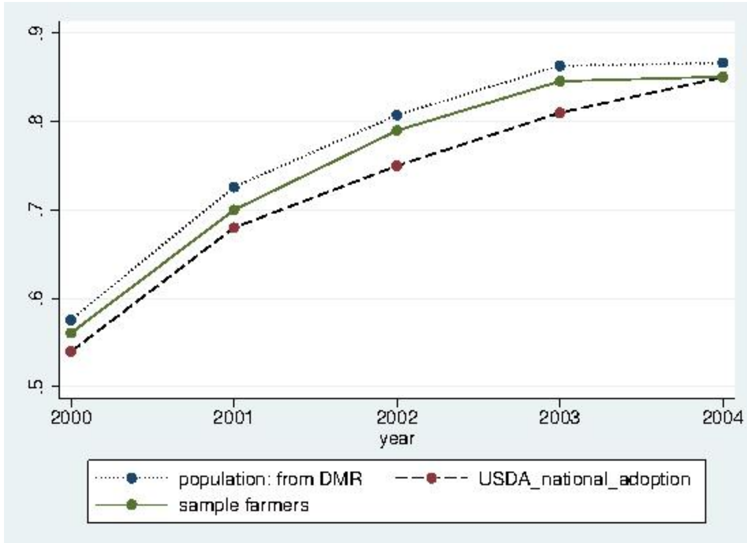

Figure 3: Average Adoption Rates: Sample vs. Population

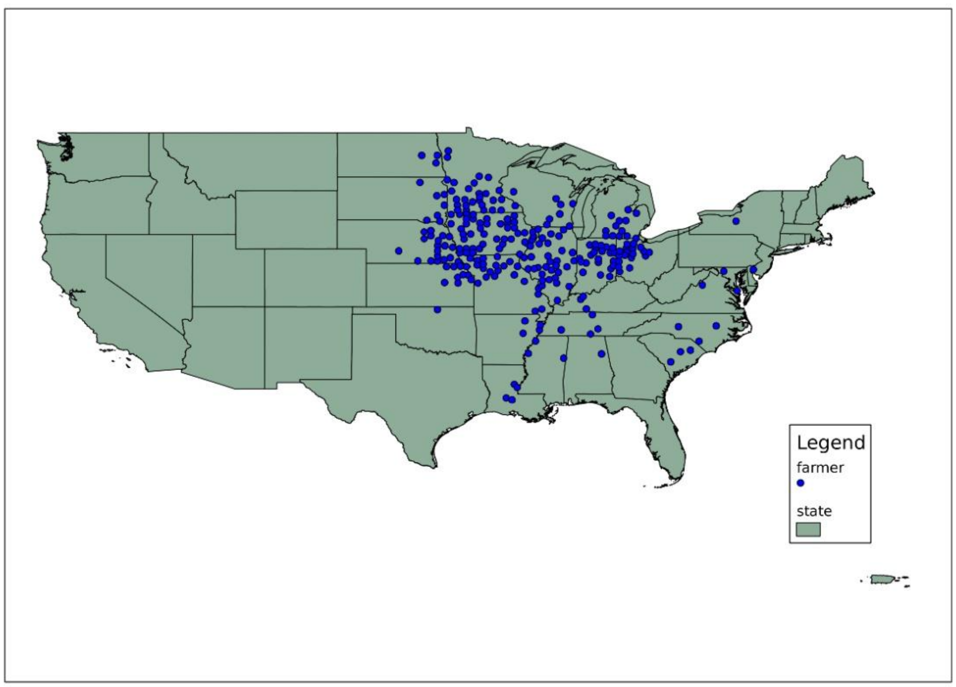

Figure 4: The Location of Selected Farmers

Since farm productivity differs by local agro-climatic conditions, farmers' experience in a southern state may not be useful to farmers located in a northern state. Therefore, we define the "neighborhood" at the Crop Report District (CRD) level, and construct the CRD adoption rate using the DMR population data. For farm size, we use the average individual soybean acreage

and the average CRD soybean acreage over years for A_i and A_{-i}. We also include the latitude and longitude of the center of the county where the sample farms are located. Table 1 shows the explanatory variables and their summary statistics. Note that on average lagged adoption rates are lower than current adoption rates, which suggests an increasing adoption pattern.

Table 1: Variable description and summary statistics

Variable	Description	Mean	SD	Min	Max
α_{it}	Farmer i 's adoption rate at time t	0.81	0.36	0.00	1.00
α_{it-1}	Farmer i 's adoption rate at time $t-1$	0.73	0.40	0.00	1.00
α_{-it}	Farmer i 's neighbors' adoption rate at time t	0.82	0.12	0.19	1.00
α_{-it-1}	Farmer i 's neighbors' adoption rate at time $t-1$	0.75	0.16	0.05	1.00
A_i	Farmer i 's total soybean acreage (acres)	328	354	40	3370
A_{-i}	Farmer i 's neighbors' total soybean acreage (acres)	100	46	41	374
P_{it}^{GM}	GM seed price paid by Farmer i at time t ($/50lb bag)	22.20	2.44	3.01	29.30
P_{-it}^{GM}	GM seed price paid by Farmer i 's neighbors at time t ($/50lb bag)	21.87	1.49	15.41	27.01
P_{it}^{Conv}	Conventional seed price paid by Farmer i at time t ($/50lb bag)	11.81	4.47	0.37	26.00
P_{-it}^{Conv}	Conventional seed price paid by Farmer i 's neighbors at time t ($/50lb bag)	11.33	4.28	0.37	25.00
Lat	Latitude of the farm	40.98	2.65	30.60	47.77
Lon	Longitude of the farm	90.77	4.90	75.35	99.82

Note: The data include 348 U.S. soybean farmers observed for five years. Only four years data are used in the estimation since we need the lagged value of the adoption rate.
Source: DMR survey data and USDA NASS website.

The average total soybean acreage for sample farmers is notably larger than that of their neighbors, perhaps because large farm owners tend to respond to the survey more often than small farm owners. However, according to Figure 3, it may not cause much attrition bias in terms of adoption patterns. Summary statistics also suggests that prices of both conventional and GM seeds paid by sample farmers are very similar to prices paid by their neighbors. Latitude and longitude records suggest that our sample farmers are concentrated in the Midwest area of the U.S.

4. Estimation

In the empirical application, we estimate both a myopic model and a forward-looking model. For both models, the simulated generalized method of moment (GMM) is used to search for the set of parameters that minimize a weighted distance between the predicted adoption path and the observed adoption path.

4.1 Myopic Model

Myopic farmers only maximize their current payoff during each time period. Thus, at time t, farmer i chooses the optimal α_{it} to maximize her current payoff u_{it} as defined in Equation (7),

$$\max_{\alpha_{it}} \quad u_{it} \quad = \left(\eta_c + \eta_{ig}\alpha_{it} - \frac{1}{2}\eta_{gc}\alpha_{it}^2 \right) A_{it} - \frac{1}{2}\rho_i A_{it}^2 \left(\alpha_{it}^2 (\varphi_{igt}^2 + \sigma_{\varepsilon}^2) + (1-\alpha_{it})^2 \sigma_{ict}^2 \right).$$

The first order condition gives

$$\alpha_{it}^*(\sigma_{igt}^2 \mid \Theta) = \frac{\eta_{ig} + A_i \rho_i \sigma_{ict}^2}{\eta_{gc} + A_i \rho_i \left(\varphi_{igt}^2 + \sigma_{\varepsilon}^2 + \sigma_{ict}^2 \right)}, \tag{10}$$

where Θ is the parameter space as defined before. And the second order condition is

$$\mu_{it}'' = -\eta_{gc} - A_i \rho_i \left(\varphi_{igt}^2 + \sigma_{\varepsilon}^2 + \sigma_{ict}^2 \right) < 0.$$

Equation (10) suggests that for any set of parameters there is a one-to-one correspondence

between the perceived GM variance φ^2_{igt} and the optimal adoption rate α^*_{it}. Since the actual adoption rate in the first period (year 2000) is known, we can obtain the perceived GM variance for year 2000 φ^2_{ig0} by solving the inverse function of $\alpha^*_{it}(\sigma^2_{igt} \mid \Theta)$, which is

$$\varphi^2_{ig0}(\alpha_{i0} \mid \Theta) = \frac{\eta_{ig} - \eta_{gc}\alpha_{i0} + A_i \rho_i \sigma^2_{ic0}}{A_i \rho_i \alpha_{i0}} - \sigma^2_{\varepsilon} - \sigma^2_{ic0}. \qquad (11)$$

We then update φ^2_{igt} for all the following years according to the Bayesian rule in Equation (3), and compute the predicted adoption rate for each farmer in all the following years according to Equation (10).

4.2 Forward-looking model

In the forward-looking model, farmers account for all future benefits when making adoption decisions. In order to compute the predicted adoption path, we make assumptions on the transition probabilities of state variables, value function of the last period and the priors of Bayesian beliefs.

Assumption on Transition Probability

Since we focus our analysis on those farmers with relatively constant soybean acreage over time, and data suggest that the average of their neighbors' total soybean acreage remains stable during the study period, we can rewrite A_{it} as A_i and A_{-it} as A_{-i}. The state variables can be reduced to $S_{it} = \{\alpha_{it-1}, \alpha_{-it-1}, z_t, A_i, A_{-i}\}$ according to the specification of φ^2_{igt} and σ^2_{ict}. Following Foster and Rosenzweig (1995) and Besley and Case (1994), we assume *Markov perfect equilibrium* for each market, which implies that farmer i and her neighbors simultaneously choose their optimal adoption rates in each time t; therefore, the pair of $\{\alpha_{it}, \alpha_{-it}\}$ forms a solution of the *equilibrium.* So the transition probability of the states is

$$P(S_{it+1} \mid S_{it}) \quad = P(\alpha_{it}, \alpha_{-it}, z_{it+1} \mid \alpha_{it-1}, \alpha_{-it-1}, z_{it}) = P(z_{t+1} \mid z_t).$$

For $P(z_{t+1} \mid z_t)$, we follow Tauchen (1986) to discretize the space of z_t to 9 equispaced points and compute their transition probabilities. See Appendix C for details.

Assumption on the Last Period

The data suggest that toward the end of the study period (year 2004), the change in the adoption rate diminishes (See Figure 3). Indeed, most farmers stop adjusting their adoption rate of the GM soybean seeds three or four years after they start their field experiment with GM seeds. Therefore, we assume that in the last period the dynamic learning process reaches the steady state, i.e., $EV_{iT+1} = EV_{iT}$ for $T \geq 5$. The Bellman equation for the last period is

$$EV_{iT} \quad = E \max_{\alpha_{iT}} \{u_{iT} + \delta EV_{iT+1}\},$$

and

$$EV_{iT+1} = EV_{iT} = \frac{1}{1-\delta} \max_{\alpha_{iT}} u_{iT} \left(S_{iT} \mid \Theta\right).$$

Based on the value of the last year, we compute the value function for all the previous years for each farmer according to the Bellman Equation.

Assumption on the Prior of Bayesian Beliefs

To update the Bayesian beliefs, we need the prior for the first period for each farmer. In the myopic case, we infer the prior belief of each farmer from their actual adoption rates in year 2000. However, in the dynamic model, the relationship between the Bayesian beliefs and farmers' adoption rates is no longer a one-to-one correspondence. If we treat the priors for each farmer as parameters as in Besley and Case (1994), it will increase the parameter space tremendously and the problem will become intractable. To overcome this problem, we use the beliefs of each farmer in year 2000 in the myopic case as reference value for the Bayesian beliefs in the

dynamic case. To account for this potential bias induced by the myopic assumption, we add a parameter b to all the myopic beliefs in 2000 and use them as the priors for the dynamic case. So the Bayesian update rule for the perceived variance of the GM profit for farmer i in 2001 follows

$$\varphi_{ig2001}^2 = \frac{1}{\frac{1}{\varphi_{ig2000}^2 + b} + \frac{G_{it}}{\sigma_\varepsilon^2} + \frac{G_{-it}}{\sigma_\varepsilon^2 + \sigma_\xi^2}},$$

where φ_{ig2000}^2 is the Bayesian belief for farmer i in 2000.

Compute the Predicted Adoption Rate

The following algorithm is used to compute the predicted adoption rate:

1. Discretize the state/control space:

The state variables are $S_{it} = \{\alpha_{it-1}, \alpha_{-it-1}, z_{it}\}$, and the control variable is α_{it}, the adoption rate. We discretize all the adoption rates $\alpha_{it}, \alpha_{it-1}, \alpha_{-it-1}$ to be 51 equal-spaced points in $[0,1]$. For the random state variable z_t, as suggested by Tauchen (1986), we discretize it into 9 equal-spaced points in an interval $[\underline{z}, \overline{z}]$, where $\overline{z} = -\underline{z} = 3\sigma^2$ and they are the lower bound and upper bound of z.

2. Simulate the random state variable z_t for each period:

We assume z_t is at its invariant state in the first period, and simulate 9 initial points according to its invariant probability. Then for each initial point we simulate a sequence for the next four years according to its transition probability.

3. Compute the Bayesian beliefs:

We compute the priors as described and update the Bayesian beliefs according to the updating rule in Equation (3).

4. Compute the value and policy functions (i.e., the optimal adoption rate under each possible state, of the last period).
5. Compute the value and policy functions for all previous years by backward induction according to the Bellman equation in (9).
6. Trace out the adoption path for each farmer based on the policy function.

4.3. Simulated GMM

Simulated GMM is used to estimate both models. For the myopic model, we solve the model for all the simulated states z_t and then take the average value. For the forward-looking model, we compute the optimal adoption path for each simulated z_t and then take the average value. In both cases, we try to find a set of parameters that minimize a weighted distance between the predicted and actual adoption rates.

Define the prediction error as $e(\theta) = \alpha_{it}^{*}(\theta) - \alpha_{it}^{s}$, where $\alpha_{it}^{*}(\theta)$ is the predicted adoption rate, α_{it}^{s} is the actual adoption rate, and let D be all the data available, i.e., $D = \{\alpha_{it}, \alpha_{-it}, A_i, A_{-i}, X_i\}$. Following Hansen and Singleton (1982), we assume that at the true parameter value θ_0,

$$E(e \mid D, \theta_0) = 0. \tag{12}$$

Then, for any function of data D, $T(D)$,

$$E(T(D)e(\theta_0)) = 0. \tag{13}$$

This fact is used to construct moments to estimate the parameters by generalized method of moments (GMM). Let k be the dimension of the parameters, l be the dimension of the moments, and $l \geq k$ due to identification requirement. Let $g_i(\theta) \equiv T_i(D)e(\theta)$, then the GMM objective function is

$$J(\theta) = n \cdot \bar{g}_n(\theta) \cdot W_n \cdot \bar{g}_n(\theta), \tag{14}$$

where $\bar{g}_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} g_i(\theta),$ and the efficient weight matrix $W_n = \left(\frac{1}{n}\sum_{i=1}^{n} \hat{g}_i \hat{g}_i{}' - \bar{g}_n \bar{g}_n{}'\right)^{-1}$, with

$\hat{g}_i = \hat{g}_i(\tilde{\theta})$ obtained from a preliminary estimation of θ with $W = I$, where I is the identity matrix. The asymptotic distribution of the estimates $\hat{\theta}$ is

$$\sqrt{n}(\hat{\theta} - \theta) \to N\left(0, (G'\Sigma G)^{-1}\right), \tag{15}$$

where $\Sigma = (E(g_i g_{i'}))^{-1}$ and $G = E\frac{\partial}{\partial \theta'} g_i(\theta)$.

5. Empirical Results

Both the myopic model and the forward-looking model are estimated. According to the discussion in Section 4.3, we chose the following instruments to facilitate the GMM estimation: a constant vector 1; last year's GM seed adoption rate of farmer i and her neighbors' adoption rate ($\alpha_{it-1}, \alpha_{-it-1}$) plus the square terms, their total soybean acreage (A_i, A_{-i}) and the square terms, farm characteristics X_i, i.e., the longitude and latitude of each county center where farms are located and the square terms, plus the GM and conventional seed prices paid by sample farmers and their neighbors (average price at CRD level)[9]. In total, the forward-looking model has 17 moments with 15 parameters and the myopic model has 17 moments with 14 parameters. Table 2 introduces the parameter definitions and corresponding initial values. The discount factor δ is set at 0.96 for the forward-looking model, following common practice in the literature (e.g. Rust, 1987; Pakes, 1986; Crawford and Shum, 2005). Part of the initial values for the myopic model is chosen based on the result from reduced form estimation as in Foster and Rosenweig (1995),[10]

[9] Although GM and Conventional seed prices are not included in our model, they are part of the factors that affect farmers' adoption behavior; therefore, price information can be used to interact with the prediction error to estimate model parameters.
[10] The analytical solution of the myopic model in equation (10) is not necessarily bounded between 0 and 1 in numerical estimation. We first run an unbounded estimation, after getting a converged parameter estimates from a continuously updated GMM estimation, we then run another two-step GMM estimation using the converged parameter values as starting values. The converged parameter values of the myopic model are also used as starting values for the forward-looking model.

and we use the estimated parameters from the myopic model as the starting value for the forward-looking model.

Table 2: Parameter definition and the initial values for the myopic model estimation

Parameter	Definition	Initial Value
η_g	Constant term of GM mean profit (+)	1.00
η_{gc}	Decreasing rate of GM mean profit with adoption	0.50
γ_0	Constant term for Conventional variance (+)	5.00
σ_ε^2	GM profit variance (+)	1.00
σ_ξ^2	Learning variance from neighbors (+)	10.00
β_0	Risk averse (+)	1.00
β_1	Farm size effect on risk averse (+)	0
λ	Parameter of AR1 process (+)	0.20
σ_v^2	Disturbance of AR1 process (+)	0.24
γ_1	Linear term for Conventional variance	1.00
$c1\_lat$	Latitude effect on mean profit	-0.19
$c2\_lat2$	Second order effect of latitude	0.08E-02
$c3\_lon$	Longitude effect on mean profit	0.25
$c4\_lon2$	Second order effect of longitude	-0.03E-02
B	Adjustment on Bayesian beliefs of GM variance	0

Myopic vs. Forward-looking

The Nelder-Mead simplex method is used to minimize the GMM objective function for both models. Table 3 presents the parameter estimates. The estimation results show that although both models share a similar structure and estimation strategy, the forward-looking model presents an adoption scenario with a subtle but important difference than what the myopic model predicts.

Table 3: Estimated parameters for myopic and forward-looking models by simulated GMM estimation

	Myopic Model		Forward-looking Model	
Parameter	**Coeff.**	**S.E.**	**Coeff.**	**S.E.**
η_g	1.28***	0.06	1.43***	0.05
η_{gc}	0.95***	0.01	0.97***	0.23
γ_0	0.97***	0.02	0.95*	0.88
σ_ε^2	0.04***	0.01E-1	0.07**	0.04
σ_ξ^2	102.03***	12.57	6.83*	4.60
β_0	2.94***	0.06	3.55***	0.29
β_1	-0.01	0.02	-0.06E-1	0.09
λ	0.34***	0.03E-1	0.39***	0.13
σ_ν^2	0.44***	0.14	0.30***	0.04
γ_1	1.09***	0.32	0.43	0.95
c1_lat	-0.17***	0.02	-0.14*	0.10
c2_lat2	0.03E-1***	0.04E-2	0. 42E-2***	0.05E-2
c3_lon	0.35***	0.15	0.38***	0.13
c4_lon2	0.01E-2	0.05E-1	-0.01E-2	0.02
B			0.06***	0.07E-2
MSE	0.077		0.068	

Note: 1. Statistical significance is denoted by *** for 1% level, ** for 5% level, and * for 10% level;
2. To compute the standard errors we used numerical derivatives with a step-size of 5% for both model.

In Table 3, the mean squared error (MSE) of the forward-looking model is smaller than that of the myopic model. Figure 5 shows the squared predicted errors from both the myopic model and the forward-looking model from 2001 to 2004. On average the squared prediction error from the forward-looking model is smaller than that from the myopic model in each year. These results suggest that the forward-looking model fits the data better and predicts better than the myopic model does, implying that soybean farmers in our sample are more likely to be

forward looking rather than myopic when they make adoption decisions. This finding is consistent with much other adoption literature (e.g. Besley and Case, 1994; Munshi, 2004).

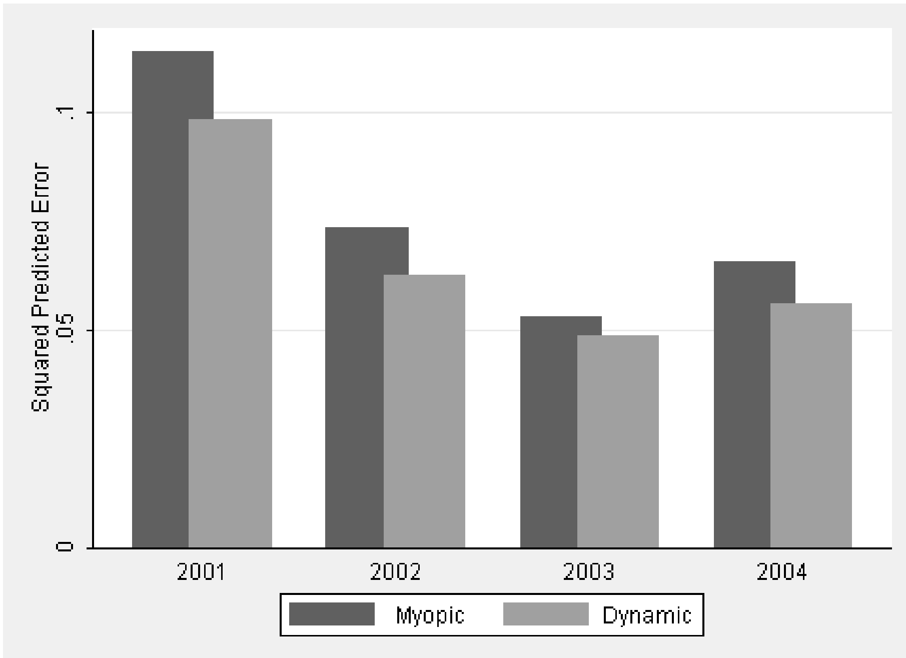

Figure 5: Squared Prediction Error: Myopic vs. Forward-looking

For those individual farms that the two models differ substantially in adoption predictions, the forward-looking model tends to perform better than the myopic model, especially in the initial period. Table 4 suggests that when the prediction difference is larger than 0.01 in squared prediction error, the forward-looking model predicts better (closer to the true adoption rate) for 89 farms in 2001, while the myopic model predicts better for only 17 farms. Indeed, our theory suggests that the myopic model underestimates the value of early adoption and, therefore, predicts lower adoption rates in early years. The forward-looking model continues to perform better than the myopic model in the following years. However, the difference in prediction power becomes smaller. Our theoretical analysis already suggests that the difference between the two models will become smaller as adoption approaches a steady state. Given the short time horizon in the finite-period game, forward-looking farmers differ less than myopic

farmer towards the end of the game, thus the decrease in prediction difference in later years is not surprising. Overall, the forward looking model predicts better in 204 observations and the myopic model predicts better in 116 observations.

Table 4: Comparison of Myopic vs. Forward-looking Model[a]

Year	Myopic Model is better	forward-looking Model is better
2001	17	89
2002	46	51
2003	31	36
2004	22	28
sum	116	204

[a] Number of observations that either myopic model or forward-looking model predicts better when the difference of their squared prediction error is larger than 0.01.

The parameter b serves as a proxy of the difference of the Bayesian belief towards the profit variance of GM seed in year 2000 between the myopic model and the forward-looking model. The estimated value of b is positive and significant. It suggests that the perceived profit risk of the GM seed in early years is higher in the forward-looking model than in the myopic model. With a higher perceived risk, learning becomes more valuable, and the potential future benefit of early adoption leads forward-looking farmers to adopt more GM soybean seed in the early years than that the myopic model predicts.

Self-learning vs. Learning from Neighbors

Based on Foster and Rosezweig (1995), we define $\rho_\varepsilon = \frac{1}{\sigma_\varepsilon^2}$ as a measure of the learning efficiency from a farmer's own experience with one unit of land, and $\rho_h = \frac{1}{\sigma_\varepsilon^2 + \sigma_\xi^2}$ as a measure of the learning efficiency from her neighbor's experience. Table 2 shows that the estimated parameter σ_ξ^2, the noise occurring in learning from neighbors, is much larger than the estimated

parameter σ_{ε}^{2}, the noise in self learning. Consequently, the learning efficiency from a farmer's own experience is much greater than the learning efficiency from neighbors during the adoption process (ρ_{ε} =14.28 vs. ρ_h =0.14). This result is consistent with findings in other related literature (e.g. Munshi, 2004; Baerenklau, 2005; Conley and Udry, 2010).

Comparing these two models, the forward-looking model identifies a stronger neighborhood effect: the estimated noise σ_{ξ}^{2} in the forward-looking model (at 6.83) is much smaller than that in the myopic model (at 102.03), therefore the forward looking model estimates a greater learning efficiency from neighbors. In this case, fitting the wrong model in empirical analysis inevitably leads to underestimating the neighborhood effects. This may be a potential explanation of the limited neighborhood effects that are found in other literature.

Mean Profit

The estimated negative slope coefficient η_{gc} in both models suggests that the marginal benefit from adopting GM seed decreases as farmers use more acreage to plant GM seed, but the net benefit from adopting GM seed ($\eta_{ig} - \frac{1}{2}\eta_{gc}$) is still positive even when farmers completely switch to GM seed. The estimated upper bound of the mean profit from adopting GM seed (η_{ig}) is higher in the forward-looking model than in the myopic model, with a slightly bigger decreasing rate of the marginal profit.

The estimated parameters with respect to farm characteristics from both models *(c1_lat, c2_lat2, c3_lon, c4_lon2)* suggest that the mean profit of GM soybean seed is higher if the farm is located in the southern and/or eastern area, but with a slightly reversed second order effect in both directions.

Other Results

The estimated parameter β_0 is positive and significant in both models, suggesting that farmers in our sample are risk averse, consistent with the literature on farmers' attitude towards new technologies. However, the effect of total soybean acreage on risk averseness is ambiguous: the estimated β_1 from both models is negative but not significant. Large soybean acreage could indicate a farmer's wealth status, and wealthy farmers tend to be less risk-averse as commonly observed in literature. On the other hand, large soybean acreage could also mean a higher switching cost, due to a long-term contract or fixed investment, which drives the adoption pattern of large farm owners to the other direction. These parameter estimates (β_0 and β_1) suggest that farmers in our sample show significant risk aversion but their risk-averseness is invariant with their farm size.

On the other hand, if farmers in our sample are indeed forward-looking, then fitting the myopic model to forward-looking farmers will generate biases in estimating the risk-averse coefficient. Table 3 shows that the coefficient β_0 estimated by the forward-looking model is larger than that estimated by the myopic model. So fitting the wrong model could underestimate farmers' risk averseness. Another noticeable difference between the estimation results of these two models is the estimated γ_1, the effect of the random state variable on the profit variance of the conventional seed. The estimated value is insignificant in the forward-looking model but positive and significant in the myopic model. Therefore, fitting the myopic model to forward-looking farmers would over-estimate the effect of the random state variable.

6. Conclusion

Besley and Case (1993) rightly state that a key factor in modeling technology adoption is "the extent to which empirical estimation is consistent with an underlying theoretical model of

optimization behavior”. In this paper, we construct and estimate two adoption models: myopic and forward-looking adoption models. We develop both the theoretic model and the empirical estimation method. Using a panel data set of 348 U.S. soybean farmers, we compare the result of the forward-looking model with the myopic model, and find that the forward-looking model fits our data better than the myopic model does, suggesting that farmers in our sample are more likely to be forward-looking. In particular, the myopic model predicts lower adoption rates in early years, implying that the myopic model fails to take account of the possible future benefits of early adoption, and therefore underestimates the value of early adoption. This finding highlights the importance of estimating an empirical adoption model consistent with underlying decision processes. It confirms that technology adoption in agriculture is likely to be a dynamic process and that farmers behave in a forward-looking manner (Griliches, 1957; Barham et al., 2004; Foster and Rosezweig, 1995).

We also find that farmers learn both from their own and their neighbors’ experiences. However, the neighborhood effect we find in our case is smaller than self-learning. GM technology in soybean seed is sensitive to individual farm characteristics; therefore, experience from one farmer may not apply to others and the true distribution of the return of GM soybean seed can only be learned by farmers’ own experiences. The myopic model, as compared to the forward-looking model, predicts even smaller neighborhood effects. If farmers are forward-looking, fitting a myopic model to them could underestimate the neighborhood effect.

Other potential biases are associated with using a myopic model on forward-looking economic agents. These include: underestimating economic agents’ risk averseness, underestimating the mean profit of the new technology, or over-estimating the random shock on the profit variance of the existing technologies. These potential biases may be good research

topics for future studies.

Recognizing farmers’ forward-looking behavior and, therefore, estimating precisely farmers’ self-learning effect and neighborhood effect are not only an important progress in the technology adoption literature but also key to delivering advanced technologies to farmers. For example, it justifies the effect of demonstration projects on surrounding farmers through information flow from experienced neighbors. It also points out that free or low cost of access to new technologies encourages farmers’ adoption. A precise estimation of these effects could be beneficial for policy makers to implement these strategies to promote adoption of new technologies.

References:

[1] U. Aldana, J. D. Foltz, B. L. Barham, and P. Useche. "Sequential Adoption of Package Technologies: the Dynamics of Stacked Trait Corn Adoption", *American Journal of Agricultural Economics*, 93(1):130–143, 2011.

[2] K. A. Baerenklau. "Toward an Understanding of Technology Adoption: Risk, Learning, and Neighborhood Effects", *Land Economics*, 81(1):1–19, 2005.

[3] B. L. Barham, J. D. Foltz, D. Jackson-Smith, and S. Moon. "The Dynamics of Agricultural Biotechnology Adoption: Lessons from RBST Use in Wisconsin, 1994-2001", *American Journal of Agricultural Economics*, 86(1):61–72, 2004.

[4] T. Besley and A. Case. "Modeling Technology Adoption in Developing Countries", *American Economic Review*, 83(2):396–402, May 1993.

[5] T. Besley and A. Case. "Diffusion As a Learning Process: Evidence from HYV Cotton", Working Papers 228, Princeton University, Woodrow Wilson School of Public and International Affairs, Research Program in Development Studies., May 1994.

[6] L. A. Cameron. "The Importance Of Learning in the Adoption of High-Yielding Variety Seeds", *American Journal of Agricultural Economics*, 81(1):pp. 83–94, 1999.

[7] T. G. Conley and C. R. Udry. "Learning About a New Technology: Pineapple in Ghana", *American Economic Review*, 100(1):35–69, March 2010.

[8] G. S. Crawford and M. Shum. "Uncertainty and Learning in Pharmaceutical Demand", *Econometrica*, 73(4):1137–1173, July 2005.

[9] J. Cummings, Ralph W. The Puebla Project. Paper prepared for meeting on social science research in rural development, Rockefeller Foundation, April 1975.

[10] G. Feder, R. E. Just, and D. Zilberman. "Adoption of Agricultural Innovations in

Developing Countries: A Survey”, *Economic Development and Cultural Change*, 33(2):255–298, 1985.

[11] A. D. Foster and M. R. Rosenzweig. “Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture”, *Journal of Political Economy*, 103(6):1176–1209, December 1995.

[12] Z. Griliches. “Hybrid Corn: An Exploration in the Economics of Technological Change”, *Econometrica*, 25(4):501–522, 1957.

[13] L. P. Hansen and K. J. Singleton. “Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models”, *Econometrica*, 50(5):pp. 1269–1286, 1982.

[14] C.-f. Huang and R. H. Litzenberger. *Foundations for Financial Economics*. North-Holland, New York, 1988.

[15] T. M. Hurley, P. D. Mitchell, and M. E. Rice. “Risk and the Value of Bt Corn”, *American Journal of Agricultural Economics*, 86(2):345–358, 05 2004.

[16] C. F. Manski. “Identification of Endogenous Social Effects: The Reflection Problem”, *The Review of Economic Studies*, 60(3):531–542, 1993.

[17] K. Munshi. “Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution”, *Journal of Development Economics*, 73(1):185–213, February 2004.

[18] A. Pakes. “Patents As Options: Some Estimates of the Value of Holding European Patent Stocks”, *Econometrica*, 54(4):pp. 755–784, 1986.

[19] J. Rust. “Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher”, *Econometrica*, 55(5):999–1033, September 1987.

[20] G. Tauchen. “Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions”, *Economics Letters*, 20(2):177–181, 1986.

[21] P. Useche, B. L. Barham, and J. D. Foltz. “Integrating Technology Traits and Producer Heterogeneity: A Mixed-Multinomial Model of Genetically Modified Corn Adoption”, *American Journal of Agricultural Economics*, 91(2):444–461, 05 2009.31

Appendix A: Bayesian Learning

A1. Self-learning

Suppose at time 0 farmer i has a prior of π_{ig0} as $N(\zeta_{ig0}, \varphi_{ig0}^2)$. If she only tries GM seed on one plot in time 0 and gets a realized profit π_{ig0}, then according to Bayesian rule, the posterior $N(\zeta_{ig1}, \varphi_{ig1}^2)$ is updated as

$$\begin{aligned} \zeta_{ig1} &= \frac{\frac{\pi_{ig0}}{\sigma_\varepsilon^2} + \frac{\zeta_{ig0}}{\varphi_{ig0}^2}}{\frac{1}{\varphi_{ig0}^2} + \frac{1}{\sigma_\varepsilon^2}} \\ \varphi_{ig1}^2 &= \frac{1}{\frac{1}{\varphi_{ig0}^2} + \frac{1}{\sigma_\varepsilon^2}}. \end{aligned}$$

If she planted GM seeds on G_0 plots at time 0 and get an average profit on each plot as $\bar{\pi}_{g0}$, then

$$\begin{aligned} \zeta_{ig1} &= \frac{\bar{\pi}_{ig0} \frac{G_0}{\sigma_\varepsilon^2} + \frac{\zeta_{ig0}}{\varphi_{ig0}^2}}{\frac{1}{\varphi_{ig0}^2} + \frac{G_0}{\sigma_\varepsilon^2}} \\ \varphi_{ig1}^2 &= \frac{1}{\frac{1}{\varphi_{ig0}^2} + \frac{G_0}{\sigma_\varepsilon^2}}. \end{aligned}$$

So the more plots this farmers tried, the more the weight of the posterior mean will goes to $\bar{\pi}_{g0}$, which converges to the true mean as the number of plots goes to infinity.

A2. Learning from neighbors

Suppose this farmer could also observe the profits of her neighbors, but with an additional noise ξ, whose variance σ_ξ^2 is assumed to be known for all farmers. Suppose her neighbors grow H_0 plots in average at time 0 and she observed an average profit as $\bar{\pi}_{ih0}$ from the neighbors, follow

the same logic of self-learning, we can rewrite the posterior as

$$\zeta_{ig1} = \frac{\bar{\pi}_{ig0}\dfrac{G_0}{\sigma_\varepsilon^2} + \bar{\pi}_{ih0}\dfrac{H_0}{\sigma_\varepsilon^2+\sigma_\xi^2} + \dfrac{\zeta_{ig0}}{\varphi_{ig0}^2}}{\dfrac{1}{\varphi_{ig0}^2} + \dfrac{G_0}{\sigma_\varepsilon^2} + \dfrac{H_0}{\sigma_\varepsilon^2+\sigma_\xi^2}}$$

$$\varphi_{ig1}^2 = \frac{1}{\dfrac{1}{\varphi_{ig0}^2} + \dfrac{G_0}{\sigma_\varepsilon^2} + \dfrac{H_0}{\sigma_\varepsilon^2+\sigma_\xi^2}}.$$

So the information from her neighborhood will accelerate the process for the posterior mean to converge to the true mean.

A3. Bayesian updating

Note that after time 0, the posterior $N(\zeta_{ig1}, \varphi_{ig1}^2)$ becomes the prior for time 1, and farmers keep updating their beliefs. So for a typical farmer, the Bayesian updating at time t is

$$\zeta_{igt+1} = \frac{\bar{\pi}_{igt}\dfrac{G_t}{\sigma_g^2} + \bar{\pi}_{iht}\dfrac{H_t}{\sigma_\varepsilon^2+\sigma_\xi^2} + \dfrac{\zeta_{igt}}{\varphi_{igt}^2}}{\dfrac{1}{\varphi_{igt}^2} + \dfrac{G_t}{\sigma_\varepsilon^2} + \dfrac{H_t}{\sigma_\varepsilon^2+\sigma_\xi^2}},$$

$$\varphi_{igt+1}^2 = \frac{1}{\dfrac{1}{\varphi_{igt}^2} + \dfrac{G_t}{\sigma_\varepsilon^2} + \dfrac{H_t}{\sigma_\varepsilon^2+\sigma_\xi^2}}.$$

Appendix B: Derivation of the mean and variance of the total profit

According to the specification in Section 2, for farmer i at time t, the perceived mean profit for GM and conventional seed from k^{th} plot are:

$$\zeta_{igt}^{k} = \eta_c + \eta_{ig} - \eta_{gc}\frac{k}{A_{it}},$$
$$\mu_{ict}^{k} = \eta_c.$$

Suppose farmer i planted a total of G_{it} plots of GM seed at time t, the mean of the total profit from GM is

$$\begin{aligned}\zeta_{igt} &= \sum_{k=1}^{G_{it}}\left(\eta_c + \eta_{ig} - \eta_{gc}\frac{k}{A_{it}}\right)\\ &= G_{it}(\eta_c + \eta_{ig}) - \frac{\eta_{gc}}{A_{it}}\cdot(1+2+\cdots+G_{it})\\ &= G_{it}(\eta_c + \eta_{ig}) - \frac{\eta_{gc}}{A_{it}}\cdot\frac{G_{it}(1+G_{it})}{2}\\ &\cong G_{it}(\eta_c + \eta_{ig}) - \frac{\eta_{gc}}{2}\cdot\frac{G_{it}\cdot G_{it}}{A_{it}}.\end{aligned}$$

The mean of the total profit from conventional seed is

$$\mu_{ict} = \sum_{k=G_{it}+1}^{A_{it}}(\eta_c) = \eta_c(A_{it} - G_{it}).$$

The total mean profit is:

$$\begin{aligned}E[\pi_{it}] &= \mu_{igt} + \mu_{ict}\\ &= G_{it}(\eta_c + \eta_{ig}) - \frac{\eta_{gc}}{2}\cdot\frac{G_{it}\cdot G_{it}}{A_{it}} + \eta_c(A_{it} - G_{it})\\ &= \eta_c A_{it} + \eta_{ig}G_{it} - \frac{\eta_{gc}}{2}\cdot\frac{G_{it}\cdot G_{it}}{A_{it}}\\ &= \left(\eta_c + \eta_{ig}\alpha_{it} - \frac{1}{2}\eta_{gc}\alpha_{it}^2\right)A_{it},\end{aligned}$$

where the optimal adoption rate α_{it} is defined as $\alpha_{it} \equiv \frac{G_{it}}{A_i}$.

Assume independence of profits from different land plots, the total variance of the profit of GM and conventional seed is:

$$\sigma^2(\pi_{it}) = G_{it}^2 \sigma^2(\pi_{igt}) + (A_{it} - G_{it})^2 \sigma_{ict}^2 = A_{it}^2 \left(\alpha_{it}^2 (\varphi_{igt}^2 + \sigma_\varepsilon^2) + (1-\alpha_{it})^2 \sigma_{ict}^2 \right).$$

Appendix C: Approximation of an AR(1) process (Tauchen 1986)

For an AR(1) process like

$$z_{t+1} = \lambda z_t + \nu_t \qquad \nu_t \sim N(0, \sigma^2),$$

Tauchen (1986) suggests an algorithm to approximate it in the following way.

1. First, discretize the space of z into equal-spaced points in an interval $[\underline{z}, \overline{z}]$, where $\underline{z} = -\overline{z}$ are the lower bound and upper bound of z. Suppose there are N points: $\underline{z} = z^1 < z^2 < \ldots < z^N = \overline{z}$.[11]

2. Suppose the length between two points is w, then the transition probability $P_{ij} = P(z^k \mid z^j)$ can be computed as

$$P_{ij} = \begin{cases} F\left(\dfrac{z^1 - \lambda z^j + w/2}{\sigma}\right), & k = 1 \\ F\left(\dfrac{z^1 - \lambda z^j + w/2}{\sigma}\right) - F\left(\dfrac{z^1 - \lambda z^j - w/2}{\sigma}\right), & 1 \le k \le N-1 \\ 1 - F\left(\dfrac{z^N - \lambda z^j + w/2}{\sigma}\right), & k = N. \end{cases}$$

3. Get the invariant probability P^z of each state.

   Given the transition probability matrix P that is computed from each P_{ij}, we can compute the

[11] Tauchen (1986) suggests that $N = 9$ is adequate for most purposes.

invariant probability of each state P^z by a contraction mapping

$$P^z = PP_0^z,$$

where P_0^z is an initial probability vector of each state.