The U. S. Corn Farmers' Genetically Modified Technology Adoption with Neighborhood Effects

Do-il Yoo

Department of Agricultural & Applied Economics

University of Wisconsin-Madison

scydl8@gmail.com

Poster prepared for presentation at the Agricultural & Applied Economics Association 2010 AAEA, CAES, & WAEA Joint Annual Meeting, Denver, Colorado, July 25-27, 2010

The U. S. Corn Farmers' Genetically Modified Technology Adoption with Neighborhood Effects

Do-il Yoo

Department of Agricultural & Applied Economics

University of Wisconsin - Madison

Introduction

Genetically Modified (GM) corn has been dramatically adopted by farmers within just a recent decade since the first generation GM varieties were commercially planted in 1996. For instance, the percentage of GM corn planted acres in 2009 amounts to 85% comparing to 25% in 2000 (USDA/ERS, 2009 [1]). Such a tremendous diffusion of GM corn in a short history comes with a question about which determinants have influenced corn farmers' adoption behaviors.

Previous literatures have analyzed the impact of farmers' characteristics on adoption behavior such as farm size, education level, risk preference, and credit access (Fernandez-Cornejo and Caswell, 2006 [2]). However, there are few empirical studies dealing with externalities of social interaction due to lack of accumulated data for GM technology. This study pays attention to neighborhood effects, which account for the tendencies that a farmer's adoption is affected by his/her neighbors' behaviors in a peer group (Brock and Durlauf, 2002 [3]). Our research object is to develop and analyze an empirical model introducing neighborhood effects.

Methodology

This study depends on discrete choice models derived from random utility theory in the sense that corn farmers choose a specific GM trait as an optimal alternative to maximize their expected profitability (utility) in a given plot and year. By traits, we categorize corn seed as herbicide-tolerance (HT), insect-resistance (IR), and their combination (stacked) seed contrast to non-GM (conventional) seed.

First, we use the multinomial logit (or conditional logit) model as a basic analysis tool assuming independence of irrelevant alternatives (IIA). Then, we adopt the concept of the mixed multinomial logit model for more realistic analysis by relaxing the IIA assumption. This results in the fact that stacked seeds share common trait properties with HT or IR seeds, so that there exists potential substitution among traits. In spite of its comprehensiveness, the mixed logit model is not so easy to be estimated due to its computational burden from simulation. Instead, we deal with the multinomial probit model by imposing normal density to parameters. Then, it can provide approximation of the mixed logit model with more convenience in running STATA.

Neighborhood Effect

Index

i Plot level for each farm household

$n_{(i)}$ The neighborhood group that plot i belongs to.

We assume $n_{(i)}$ as a Crop Report District (CRD) in a year.

l Possible alternative (choice/ GM trait)

<simpler version>		< More specified version>
$l=0$: Conventional (non-GM) seed		$l=0$: Conventional (non-GM) seed
$l=1$: HT-only seed		$l=1$: HT-only seed
$l=2$: IR-only seed		$l=2$: IR-only-RW (Root worm) seed
$l=3$: STACK seed	*or*	$l=3$: IR-only-CB (Corn Borer) seed
		$l=4$: Double Stacked Seed
		$l=5$: Triple Stacked Seed
		$l=6$: Quadraple Stacked Seed

Neighborhood Effects

$$NE_i = J_{n_{(i)},l} \cdot \left(\frac{1}{I_{n_{(i)}} - 1} \cdot \sum_{j \neq i, j \in n_{(i)}}^{I_{n(i)}} \mathbf{1}\{w_i = l\} \right)$$

, where NE_i indicates "Neighborhood Effects" for plot i

$J_{n_{(i)},l}$ A measure of the strength of social utility for choice l at group $n_{(i)}$

We use acre share of $l-$trait is planted for each group $n_{(i)}$ as a proxy.

$I_{n_{(i)}}$ Total number of plots in group $n_{(i)}$

w_i Choice for plot i

$\mathbf{1}\{\cdot\}$ Indicator function

Here, neighborhood effects are defined as a plot's expectation of the percentage of plots which makes the same choice in a peer group

Model Specification

Plot-level Utiity (Profitability)

$$U_{i,l} = \beta' Z_{i,l} + \varepsilon_{i,l}$$

, where $Z_{i,l}$ is a deterministic (observable) component.

$Z_{i,l}$ is composed as the following factors:

$$Z_{i,l} = \left(\{ farmsize_i \}, \{ seed\ price_{i,l}, NE_{i,l} \}, \{ HHI_{n_{(i)}}, year_{n_{(i)}} \} \right)^T$$

The first, second, and third partition are $i-$, $l-$, and $n_{(i)}-$specific factor, respectively.

$farmsize_i$	Farm size
$seed\ price_{i,l}$	Paid seed price for $l-$trait corn seed at plot i
$NE_{i,l}$	Neighborhood Effects for plot i when its choice is l
$HHI_{n_{(i)}}$	Herfindahl Index of GM trait companies at group $n_{(i)}$
$year_{n(i)}$	Time variable

, where $\varepsilon_{i,l}$ is a latent (unobservable) component.

Multinomial Logit Model: $\varepsilon_{i,l}$~i.i.d. & extreme value distribution

Multinomial Probit Model: $\varepsilon_{i,l}$ ~ Normal distribution & correlated

Choice Probability & Likelihood Function

Multinomial Logit Model:

$$\Pr(\omega_i = l) = \frac{\exp(\beta' Z_{i,l})}{\sum_k \exp(\beta' Z_{i,k})}, \qquad L = \prod_i \prod_l \Pr(\omega_i = l)$$

Mixed Multinomial Logit Model:

For simplicity, we assume a random intercept model for each $l-$choice.

Choice probability is rewritten as $U_{i,l} = (\beta_{i,l}, \beta^1, \beta^2, ..., \beta^5)(\mathbf{1}_{const}, Z^1_{i,l}, Z^2_{i,l}, ..., Z^5_{i,l})^T$

Note that we have 5 explanatory variables as is described above.

$$\Pr(\omega_i = l) = \int \frac{\exp(\beta_{i,l}' Z_{i,l})}{\sum_k \exp(\beta_{i,k}' Z_{i,k})} \cdot f(\beta) d\beta,$$ where $f(\beta)$ is a p.d.f. for β.

We assume β follows nomial distribution for tractability.

Then, we can utilize multinomial probit as an approximation of mixed logit , which requires burdensome simulation procedure.

Data

This study relies on a unique panel data set about the U.S. corn seed market over the period 2000 ~ 2007 across the United States, which includes farm level information about corn seed prices, planted acreage, and seed types by traits: conventional, HT, IR, and other kinds of stacked GM seeds. In order to avoid periodical and regional bias, a crop reporting district (CRD) for each year is assumed to correspond to a neighborhood group under the same agro-climatic conditions.

Descriptive Statistics

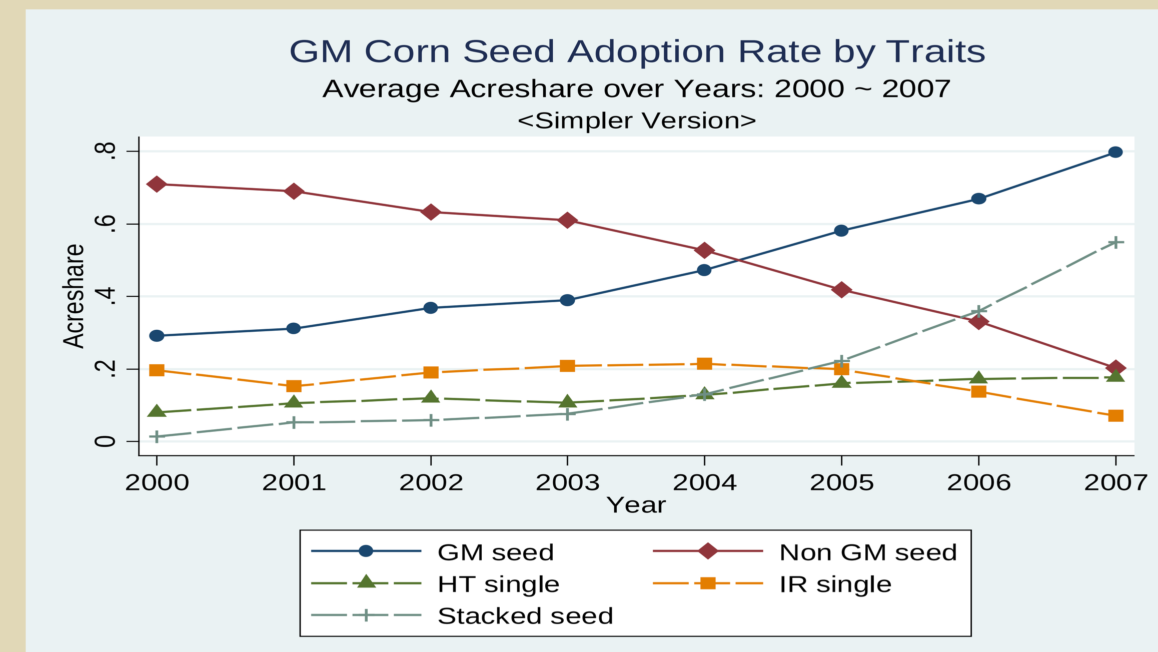
Simpler Version			Specific Version		
l	Trait	%	l	Trait	%
0	CONV	50.59	0	CONV	50.59
1	HTonly	13.89	1	HTonly	13.89
2	IRonly	16.20	2	IRonly_RW	0.65
3	STACK	19.32	3	IRonly_CB	15.55
			4	DOUBLE	14.21
			5	TRIPLE	4.64
			6	QUADRAPLE	0.47

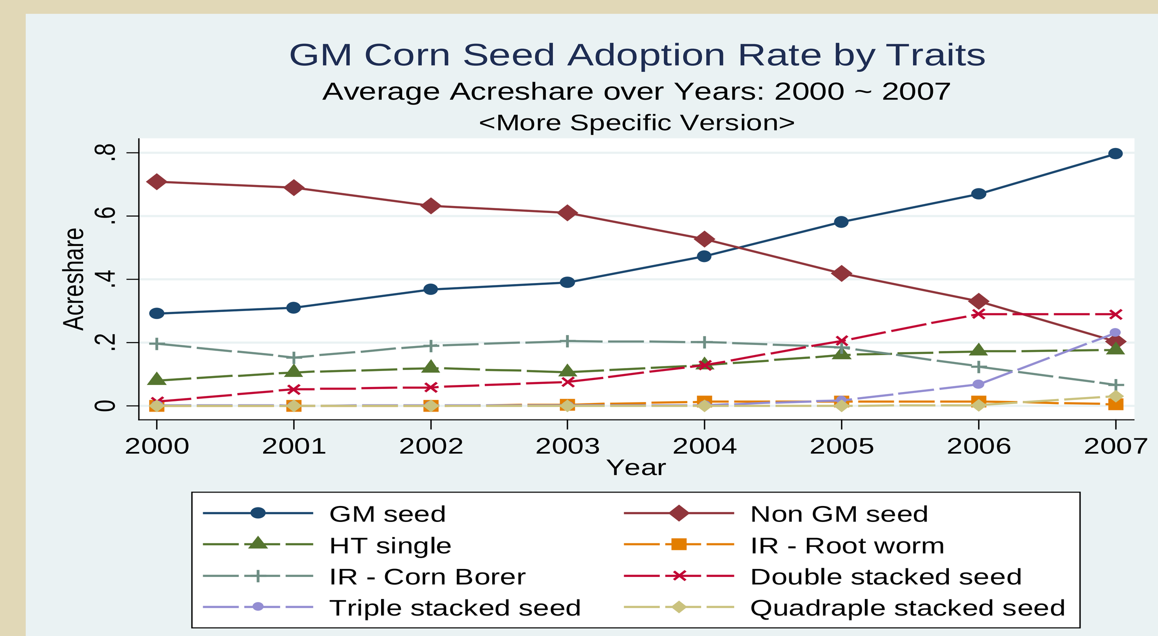
Explanatory Variable	Description	Mean	Std. Dev.
farmsize	Farm size	612.6497	743.3472
pd_S1	Price Difference b/w l – trait and CONV seed (simpler)1/	52.98297	36.32256
pd_S2	Price Difference b/w l – trait and CONV seed (specific)1/	52.98297	36.52038
NE	Neighborhood Effect Term b/w l – trait and CONV seed 2/	0.0909552	0.2249764
HHI	Herfindahl Index of trait companies	0.2334269	0.2220434

Note:
The number of observation is 168766 (87 missing values among 168853 observations).
The number of possible group, defined by each CRD per year, is 1890.
1/, 2/: These variables are adjusted for the multinomial logit model in STATA, so that the difference between base choice (CONV) is introduced into the model.

Adoption Trends

The diffusion of GM corn seed indicates two features: 1) adoption rate for GM seed increase rapidly from 2000 through 2007. 2) the trend is switching from single trait seed to stacked seed.





Results

Multinomial Logit Model

Regression is done by 'mlogit' using STATA 10.1. As a bench mark, we run the multinomial logit model both for simper and for specific version. Here, we shows more specific version.

Choice	Variables	Coef.	Std. Err.	Choice	Variables	Coef.	Std. Err.
1: HTonly	*farmsize*	0.0001345	0.000024	4: DOUBLE	*farmsize*	0.0003318	0.0000223
	pd_S2	-0.1769531	0.0010975		*pd_S2*	-0.1496967	0.0010762
	NE	8.472343	0.1962254		*NE*	9.793068	0.1936842
	HHI	-0.6651991	0.1530203		*HHI*	-2.750114	0.1506423
	year	0.2548347	0.0101827		*year*	0.4966211	0.0102088
	_cons	6.363085	0.061201		*_cons*	4.909983	0.618006
2: IRonly_RW	*farmsize*	0.0002856	0.0000447	5: TRIPLE	*farmsize*	0.0005617	0.0000262
	pd_S2	-0.1416524	0.0018599		*pd_S2*	-0.148208	0.0012117
	NE	6.857784	0.3230526		*NE*	-8.459649	0.2836903
	HHI	-3.424193	0.3769859		*HHI*	7.558296	0.1471589
	year	0.7295871	0.231395		*year*	1.676857	0.249841
	_cons	1.304541	0.1255832		*_cons*	-5.177418	0.1505115
3: IRonly_CB	*farmsize*	0.0002944	0.0000225	6: QUADRAPLE	*farmsize*	0.0005637	0.0000392
	pd_S2	-0.1673809	0.0010784		*pd_S2*	-0.1437709	0.0020467
	NE	6.663236	0.1932355		*NE*	-8.080006	0.6160253
	HHI	-0.4944043	0.1474191		*HHI*	7.114547	0.249042
	year	0.143036	0.0098224		*year*	3.286272	0.1592169
	_cons	6.849262	0.0597514		*_cons*	-18.4625	1.091314

Note:
The base alternative is l = 0 (CONV).
Number of obs: 168766, Log Likelihood: -101498.58
All estimates are statistically significant at 1% level.

Comparison among Models – for Simpler Version

Due to computational burden, we apply each model to the simpler version only. Using STATA, we use 'mlogit' and 'mprobit' for the multinomial logit and multinomial probit model, respectively.

Choice	Variables	Multinomial Logit	Multinomial Probit
1: HTonly	*farmsize*	0.0001218	0.0000606
		(0.0000236)	(0.0000148)
	pd_S1	-0.1740552	-0.1018367
		(0.0010742)	(0.000507)
	NE	7.025023	4.60956
		(0.1788384)	(0.1046804)
	HHI	1.270094	0.7648978
		(0.132381)	(0.083231)
	year	0.2146882	0.1061755
		(0.0099121)	(0.0060487)
	_cons	5.998705	3.24636
		(0.0576056)	(0.0301173)
2: IRonly	*farmsize*	0.0002811	0.0001684
		(0.0000219)	(0.0000139)
	pd_S1	-0.1631701	-0.0957957
		(0.0010478)	(0.0004735)
	NE	5.506584	3.48357
		(0.1774311)	(0.1040001)
	HHI	0.9360481	0.5045319
		(0.1288071)	(0.0816099)
	year	0.1226062	0.0305824
		(0.0095273)	(0.0057711)
	_cons	6.528903	3.715119
		(0.056255)	(0.0292489)
3: STACK	*farmsize*	0.0003442	0.0002185
		(0.0000214)	(0.0000134)
	pd_S1	-0.1416715	-0.0819779
		(0.0010146)	(0.000449)
	NE	4.423635	2.710534
		(0.1706461)	(0.098722)
	HHI	3.423928	2.452767
		(0.112435)	(0.0694432)
	year	0.5732241	0.3653961
		(0.0097598)	(0.0059218)
	_cons	3.498774	1.504215
		(0.0571771)	(0.0293186)
Log Likelihood		-91304.839	-93725.397
Number of Obs.		168766	168766

Note:
The base alternative is l = 0 (CONV).
The values in parenthesis indicate standard deviation.
All estimates are statistically significant at 1% level.

Conclusions

According to the results, all the coefficients have expected sign (e. g., negative in terms of price factor, and positive sign with respect to neighborhood effects), and statistically significant. Neighborhood effects affects the log odds of HTonly to non GM seed most, and that of STACK least. HHI influences the probabilities of choosing STACK seed most. Also, when IIA is assumed, the multinomial logit model overestimates parameters than the multinomial probit.

This study makes contribution to introducing neighborhood effects into adoption research. Future work may be put on dynamics of learning based on interaction over years.

References

[1] U. S. Department of Agriculture/Economic Research Service, July 2009, "Adoption of Genetically Engineered Crops in the U. S.: Corn Varieties," USDA, ERS, Washington, D. C., from http://www.ers.usda.gov/data/BiotechCrops/ExtentofAdoptionTable1.htm

[2] Fernandez-Cornejo, J. and M. Caswell, April 2006, "The First Decade of Genetically Engineered Crops in the United States," Economic Information Bulletin No. 11., USDA, ERS, Washington D. C.

[3] Brock, W. and S. Durlauf, 2002, "A Multinomial Choice Model with Neighborhood Effects," *American Economic Review*, 92: 298-303.