Brand Inertia in Seed Demand: State Dependence and Heterogeneity

Jinjing Luo, GianCarlo Moschini, and Edward D. Perry *

This draft: Nov 26, 2019

Abstract:

A commonly observed feature of differentiated product markets is brand inertia, the tendency of consumers to purchase brands they have purchased in the past. In this paper, we develop and estimate a micro-level random coefficients logit model to study two competing explanations of brand inertia, state dependence and heterogeneity, in the U.S. soybean seed industry. Specifically, heterogeneity is captured by brand-specific random coefficients and state dependence is incorporated through a brand purchase history variable. We further deal with two important issues in the identification: we apply a correction to the initial conditions problem similar to the procedure outlined in Wooldridge (2005); and to deal with price endogeneity, we use the control function approach in our nonlinear regression environment. The model is estimated using a large dataset of more than 200,000 seed purchase decisions by roughly 28,000 farmers over the period 1996-2016. We find that state dependence and heterogeneity are both important features of seed demand. On average, farmers are willing to pay (WTP) an additional $6.77/unit for a brand if it was purchased in the previous period, equivalent to about 15% of the average retail price. We also find that farmers are willing to pay large premiums for brand labels and the glyphosate tolerance (GT) technology, however there is considerable heterogeneity in this willingness. To investigate the implications of state dependence, especially as it relates to the introduction and diffusion of the GT innovation, we simulate several counterfactual scenarios (with/without state dependence and/or the GT technology). Our simulations show that state dependence has little effect on the diffusion of the GT technology, but it functions as a cushion for the structural benefits brought about by the innovation—it reduces the gains/losses in brands' market shares. We also show that there is an "early adoption" advantage associated with the marketing of the GT trait, which is reinforced by state dependence.

Key Words: Brand inertia, Heterogeneity, Mixed logit, State dependence, Seed demand, Technology innovation.

* Jinjing Luo is a PhD candidate and GianCarlo Moschini is a professor and Pioneer Chair, Department of Economics, Iowa State University. Edward D. Perry is an assistant professor, Department of Agricultural Economics, Kansas State University.

Contact: Jinjing Luo, lexiluo@iastate.edu.

1. Introduction

The extent to which brand loyalty matters for demand has long been a motive of interest in economics and marketing (Bronnenberg, Dubé, and Moorthy 2019). In differentiated product markets, a well-established empirical regularity is brand inertia: individuals are more likely to purchase a brand if they have purchased it in the past. Among the potential behavioral explanations for this tendency, researchers have been particularly interested in the importance of *state dependence*, defined as the causal dependency of an individual's future choices on their current state (Heckman 1981; Dubé, Hitsch and Rossi 2010). A growing body of research has shown that the presence of state dependence can have important implications for the extent of market power and pricing behavior (Dubé, Hitsch and Rossi 2010), market structure (Dubé, Hitsch and Rossi 2009), the price effects of mergers (MacKay and Remer 2019), and the persistence of brand shares (Bronnenberg, Dhar, and Dubé 2009; Bronnenberg, Dubé, and Gentzkow 2012). Whereas the extant literature has focused mainly on the consumer-packaged goods (CPG) industry,[1] this paper investigates brand inertia and state dependence in the context of an important agricultural input market: the U.S. soybean seed industry.

Over the last few decades, the seed industry has been characterized by considerable growth and consolidation (OECD 2018). Much of this has been driven by the development and rapid diffusion of genetically engineered (GE) crops. First introduced in the mid-1990s, GE varieties embedding herbicide tolerance and/or insect-resistance provided farmers with drastically new technological solutions for weed and pest management. As a result, GE crops were met with considerable success and now exceed 90% of planted U.S. acreage in corn, soybeans, and cotton (Barrow, Sexton, and Zilberman 2014). The commercialization of GE varieties required access to both GE traits and elite germplasm, the latter arising from decades of traditional breeding efforts. Whereas GE traits were overwhelmingly developed by one company

[1] In addition to those mentioned above, other studies in the CPG industry include Keane (1997), Seetharaman and Chintagunta (1998), Seetharaman (2004), and Horsky, Misra, and Nelson (2006). Sudhir and Yang (2014) and Train and Winston (2007) study the automobile industry, and Handel (2013) analyzes the health insurance industry.

(Monsanto), the ownership of germplasm was more dispersed. The highly complementary nature of these two building blocks (Graff, Rausser, and Small 2003) led to an early wave of acquisitions and mergers (Fernandez-Cornejo, 2004). Furthermore, the diffusion of GE crops was facilitated by Monsanto's aggressive licensing of GE traits to other seed suppliers, a contractual strategy that also benefited from a parallel major strengthening of intellectual property for plants (Clancy and Moschini, 2017).

Among the major U.S. crops, the soybean seed industry has perhaps undergone the largest transformation. The once common farming practices of saving harvested soybeans for seed use, and/or purchasing publicly developed varieties, have been replaced by the almost complete reliance on new proprietary commercial soybean varieties that embed the GE trait for glyphosate tolerance (GT).[2] An ongoing area of research has sought to assess the implications of these changes for the industry and the welfare of its main players: trait developers, seed companies, and farmers. An essential ingredient for this research program is the estimation of seed demand. Beyond the assessment of the value of product innovation (Ciliberto, Moschini, and Perry 2019), a suitable seed demand model would permit the investigation of other questions of interest, including the exercise of market power, related antitrust concerns that may arise, and the role of brand loyalty.

In this paper, we develop and estimate a micro-level structural model of U.S. soybean seed demand. Specifically, we estimate a random coefficients logit model that allows for the presence of state dependence in farmers' preferences for brand labels. To estimate demand, we draw on a dataset containing more than 200,000 seed purchase decisions by roughly 28,000 U.S. soybean farmers during the 1996-2016 period. These unique data provide the requisite information on seed purchase histories, seed characteristics, and prices. In developing and estimating the model, our main objectives are to: (i) identify the dollar value of state dependence for brand labels in the soybean seed industry; (ii) investigate whether farmer heterogeneity is an important feature of the demand for brand labels and GE glyphosate tolerance; (iii) assess the implications of state dependence for the adoption rate

[2] Consider, for example, that in 1970 about 70% of planted soybeans were public varieties (Fernandez-Cornejo, 2004). Based on the data used in this paper, by 2016 this fraction is less than 1%.

of a new major product innovation (in this case, the GT trait); and (iv) assess whether the presence state dependence confers an advantage to early providers of a new technology (the GT trait).

The model we develop and estimate must address two important issues. The first issue concerns the identification of state dependence. The basic problem is that brand persistence or inertia (sometimes referred to as stickiness) can arise because of genuine state dependence *or* because of heterogeneity (Heckman 1981; Keane 1997). Heterogeneity describes the fact that individuals may simply have different, state-invariant, preferences for a brand. Failure to properly control for heterogeneity will tend to exaggerate the presence of state dependence. A related but distinct issue is the *initial conditions problem* (Heckman 1987; Arulampalam and Stewart 2009; Akay 2012; Simonov et al. 2019). This problem arises when the researcher does not observe an individual's entire purchase history and, if not properly accounted for, will also tend to exaggerate the extent of state dependence.

To control for heterogeneity, we permit farmers to have normally distributed preferences for all brands. To address the initial conditions problem, we apply a correction similar to the procedure outlined in Wooldridge (2005). In particular, we include brand-specific indicator variables that code for whether an individual purchased that brand in their first period of observation. Despite such control for heterogeneity, it is still possible to obtain spurious state dependence if the assumed distribution for heterogeneity deviates significantly from the true distribution. Thus, as a final check for whether we have identified genuine state dependence, we conduct a reshuffling procedure similar to Dubé, Hitsch and Rossi (2010). The basic idea of this procedure is to reshuffle each individual's choice sequence in a random way and then re-estimate the model. If structural state dependence remains, then this suggests that unobserved heterogeneity has not been sufficiently accounted for.

The second issue we face is the well-known problem of price endogeneity in demand models of differentiated products (Berry, 1994; Berry, Levinsohn, and Pakes 1995; Nevo 2000). Although our model is estimated using individual choices, which may alleviate some concerns about endogeneity (Goldberg 1995), there may still remain certain unobservable factors correlated with both the price and demand. The most common solution to this problem is to use two stage least squares (2sls) with instrumental variables (IVs). This approach, however, cannot be directly applied in non-linear individual-level discrete choice models (Train 2009). Therefore, we implement a control function approach, as outlined in Petrin and Train (2010) and Wooldridge (2015). Much like 2SLS, this consists of running a first-stage regression of price on all model variables and a set of excluded IVs. We then compute the predicted residuals from this first-stage regression and include them as a control variable in the random coefficients logit model. For IVs, we use the previous year's soybean futures price interacted with brand and GT trait dummies. These IVs are in the spirit of the cost-brand interaction IVs used by Berto Villas-Boas (2007), and they exploit the fact that the previous year's futures price affects a seed firm's production costs.[3]

Overall, we find significant evidence of structural state dependence, even after controlling for persistent unobserved farm-level heterogeneity. On average, having a previous experience with a brand increases its value by about \$6.77/unit of soybeans, equivalent to about 15% of the average price of \$45/unit. This estimate is quite similar to the dollar value of brand loyalty estimated for orange juice in Dubé, Hitsch and Rossi (2010). We also find that, on average, farmers are willing to pay large premiums for brand labels and for GE traits, although this willingness to pay (WTP) can vary widely across farmers. For example, from 2011-2016, farmers' mean WTP for the GT trait was \$22.89/unit, with more 15% of farmers valuing GT at \$35 or more and another 15% of farmers valuing it at \$10 or less.

Using the model estimates, we assess some potential implications of state dependence for the introduction and diffusion of the GT technology. We first consider whether structural state dependence impacted the diffusion of GT soybeans. Intuitively, state dependence could impact the diffusion rate of a new product attribute if, in the early stages of adoption, the presence of state dependence significantly increased the likelihood of new adopters becoming repeat or permanent adopters. To investigate this possibility, we simulate GT

[3] Seed firms contract with individual farmers to grow their future commercial seed supply and this is usually done in the region where the seed will eventually be sold (Lamkey 2004). A farmer's opportunity cost of growing seed for a company is what they could have obtained on the market. Thus, a higher futures price will tend increase production costs for seed firms. Note that by interacting the futures price with GT and brand dummies we aim to capture product-specific cost impacts.

adoption rates under a scenario with state dependence and under a scenario without state dependence. Overall, we find little evidence of inertia impacting the rate of GT adoption. We attribute this, in part, to the fact that most brands in the industry added GT to their varieties very quickly (consistent with Monsanto's strategy to aggressively license the GT trait to other seed firms).

A second counterfactual we consider concerns whether state dependence led to an "early mover" advantage for brands that were the first to embed and offer the GT trait in their seed varieties. In certain ways, this question is similar in nature to the question of whether there is an advantage to be the first brand to locate in a particular geographic region (Bronnenberg, Dhar, and Dubé 2009; Bronnenberg, Dubé, and Gentzkow 2012). To assess this issue, we used the estimated model to simulate market shares in four settings: (i) an environment with both the GT trait and state dependence; (ii) an industry without the GT trait but still with state dependence; (iii) an industry with the GT trait but without state dependence; and (iv) an industry with neither GT nor state dependence. Comparing (i) and (ii) reveals the impact of GT on brand shares when state dependence is present, and comparing (iii) and (iv) reveal the impact of GT on brand shares when state dependence is not presence.

The analysis and results provided in this paper contribute to the literature in several ways. First, we provide new evidence on the *distribution* or heterogeneity of U.S. farmers WTP for the major U.S. soybean seed brands, and for the GT trait, over time. This complements the results reported by Ciliberto, Moschini, and Perry (2019), whose WTP estimates for GE traits were not modeled to vary across farmers. The second contribution of this study is to provide dollar value estimates of brand loyalty in an important U.S. agricultural input market. While brand loyalty in related industries has been documented using survey and interview-based evidence (Kohls et. al., 1957; Funk and Vincent, 1978; Kool, 1994; Harbor, Martin, and Akridge, 2008; Sellars and Gunderson, 2018), there are no studies of brand inertia using revealed preference data in an agricultural context. Finally, to our knowledge, no other study has considered whether state dependence can confer an "early-mover" advantage in the context of an innovation in product attributes. Whereas research in Bronnenberg, Dhar, and Dubé (2009) and Bronnenberg, Dubé, and Gentzkow (2012) find there is a large, persistent advantage to being the first brand in a particular *geographical* location, our exercise assesses whether state dependence confers an advantage to those brands first to locate in a new dimension of the *product space*.

The rest of this paper is organized as follows. In Section 2, we provide background information on the U.S. soybean seed market. Section 3 presents the data used in the econometric regression. In Section 4, we develop the demand model, discuss the identification strategy, and present the estimation process. Section 5 presents the estimation results, followed by the implied WTP distributions and demand elasticities. Using simulation, some implications of state dependence are considered in Section 6. Section 7 concludes.

2. Soybean Seed Industry Background

The U.S. seed industry has grown considerably over the last few decades, fostered by sustained demand domestically and abroad.[4] Part of this growth has been driven by technological innovation, the result of significant research and development (R&D) investments, partly owing to the changing landscape of intellectual property rights. As the industry has grown it has also experienced considerable consolidation and rising seed prices (Fernadez-Cornejo 2004). A major development affecting seed markets, maize and soybeans in particular, has been the introduction of genetically engineered (GE) traits in the mid-1990s. By using breakthrough recombinant DNA techniques of modern biology, it became possible to integrate certain foreign genes (from bacteria) into the germplasm of elite crop varieties. These genes confer traits to the resulting "transgenic" crops, such as herbicide tolerance and insect resistance, which are highly valued by growers (Moschini 2008).

The GE revolution in the seed industry also benefited from the general strengthening of intellectual property rights for biological innovations (Moschini 2010). This is particularly important for soybean seeds, the focus of this paper. Soybean varieties are self-pollinating, meaning that they reproduce true to type (unlike hybrid maize, for example). Thus, prior to

[4] The size of the global commercial seed market was estimated at about USD 12 billion (ISAAA, 2016) in the United States and around USD 52 billion worldwide (Syngenta, 2016) in 2014.

the advent of GE varieties, farmers could rely on saved seeds (from the previous harvest) and could access, essentially at cost, varieties developed and released by public institutions (state universities). The introduction of patented GE traits, and the associated increased use of trade secrets and contracts, effectively permitted the industry to develop proprietary seed products (Clancy and Moschini 2017). This greatly increased the profitability of R&D in plant breeding, which lead to increased investments and an early wave of industry consolidation through mergers and acquisitions (Fernandez-Cornejo 2004). By the year 2000, the two largest firms (Monsanto and Dupont) accounted for about 40% of the U.S. soybean seed market, a combined share that has risen to about 60% in recent years.

Soybeans constitute the second most planted crop (after maize) in the United States. Unlike genetically-engineered corn varieties, which can have several traits—glyphosate tolerance (GT), corn borer resistance, root worm resistance, and their combinations—the only trait with major commercial relevance during our study period (1996-2016) has been glyphosate tolerance.[5] Glyphosate is a powerful, broad-spectrum herbicide used in combination with GT crops. It can kill approximately 99% of non-glyphosate resistant weeds without harming GT varieties (Wechsler, McFadden, and Smith, 2018). By reducing the need to use tillage, as well as multiple types of herbicides, GT varieties permit an extremely effective (and simplified) weed control strategy (Perry, Moschini, Hennessy, 2016). Because of this, GT soybeans were rapidly adopted: first commercially introduced in 1996, GT varieties accounted for more than 50% of the market by 1999, and more than 90% by 2007. This is despite the fact that GT soybeans command a significant price premium (Schenkelaars et al., 2011; OECD, 2018). Indeed, previous resarch has found that U.S. farmers' willingness to pay (WTP) for the GT trait far exceeds its cost, resulting in significant net economic gains, especially in recent years (Ciliberto, Moschini, and Perry, 2019; Shi, Chavas, and Stiegert, 2010; Fernandez-Cornejo, Hendricks, and Mishra, 2005).

[5] GE varieties tolerant to glufosinate did not achieve commercial relevance till very recently and, as in Ciliberto, Moschini, and Perry (2019), we do not distinguish between conventional and glufosinate tolerant varieties in our empirical analysis.

The marketing of seed varieties relies heavily on long-standing and well-known brands such as Pioneer and Asgrow. In addition, several brands can be marketed by the same parent company. For example, the company Dupont has primarily sold varieties under the Pioneer brand, whereas Monsanto has marketed varieties under several brands such as Asgrow, DeKalb, and Channel. Each brand typically offers multiple distinct varieties that differ in characteristics such as glyphosate tolerance, soybean cyst nematode resistance, relative maturity, and tolerance to iron deficiency chlorosis. Most brands currently market both conventional and GT varieties.[6]

Table 1. Brand Market shares

Market share	1996-98	1999-2004	2005-10	2011-16
Monsanto				
Asgrow	0.133	0.166	0.161	0.217
Channel	0	0	0.006	0.032
DeKalb	0.078	0.054	0.037	0.003
Kruger	0.013	0.017	0.013	0.006
DuPont				
Pioneer	0.186	0.214	0.268	0.285
Syngenta				
Golden	0.037	0.061	0.027	0.001
NK	0.051	0.054	0.087	0.079
Dow				
Mycogen	0.024	0.021	0.013	0.020
Public				
Public	0.063	0.026	0.007	0.005
Others				
Beck's	0.007	0.013	0.016	0.029
Croplan	0.012	0.025	0.033	0.030
Growmark	0.018	0.014	0.011	0.008
Stine	0.041	0.031	0.024	0.021

[6] Two exceptions are Channel, a Monsanto brand, which entered the market in 2010 and only offers GT varieties, and public providers (mainly state university programs) who offer only conventional varieties (See Fernandez-Cornejo (2004), p. 36, for a list of major public breeders).

Average market shares for the 13 largest brands over the considered timespan are reported in Table 1. Brands are grouped by their well-known parent companies such as Monsanto, DuPont, Syngenta, and Dow AgroSciences.[7] We separately categorize public seeds and the seeds sold by brands not owned by the big four parent companies.[8] Table 1 also illustrates some turnover in brands. Golden Harvest was phased out by Syngenta is 2012 and, in recent years, DeKalb is also being phased out (Monsanto is focusing this brand mostly in the maize seed market).

3. Data

The data used in this paper pertain to seed purchases by a large and representative sample of U.S. soybean farmers. This data is drawn from a proprietary dataset assembled by Kynetec USA, a market research company that specializes in the collection of survey data in U.S. agriculture. The data span the 21-year period 1996-2016. For each year, the seed purchases of more than 3,500 soybean farmers are recorded. The sample itself is constructed to be representative at the crop reporting district (CRD) level.[9] Each soybean farmer in the sample is observed to make one or more seed purchases and the data contains detailed information on the nature of the purchase and the variety (e.g., variety name, brand, parent company, GE traits, price, amount of seed, acres planted). Although this is not a balanced panel data set, a large portion of farmers are observed over multiple years (and multiple purchases are

---

[7] The company names in Table 1 reflect the industry configuration as of 2016, the last year of our data. Since then, major mergers and acquisitions are re-shaping the ownership structure of the industry—the acquisition by ChemChina of Syngenta in Apr 2017, the merger of Dow and DuPont in Sep 2017, and the acquisition by Bayer of Monsanto in June 2018. The agricultural concerns consolidated by the Dow-Dupont mergers were subsequently spun off as Corteva in 2019.

[8] The ownership of each brand as reported in Table 1 also pertains to 2016, the last year of our data. Brands' affiliation with their parent company in some cases was the result of market consolidation that took place earlier in our sample. This is particularly true for Monsanto, who acquired Asgrow in 1997, DeKalb in 1998, Channel in 2004, and Kruger in 2006. Also, DuPont acquired Pioneer in 1999; Syngenta acquired NK in 2000 and Golden Harvest in 2004; Dow acquired Mycogen in 1998.

[9] CRDs are regions identified by National Agricultural Statistics Service of the U.S. department of Agriculture (USDA). Each U.S. state comprises several CRDs, and each CRD includes multiple counties.

observed in the same year). As discussed further below, multiple observations per farmer are essential for identifying/disentangling the elements of state dependence and time-invariant unobserved heterogeneity as drivers of farmers' purchase decisions.

3.1. Products

As noted, in this paper we develop and estimate a farm-level discrete choice model of soybean seed demand. An essential ingredient for this model is the definition of a "product." The product definition (or product space) partially defines farmers' choice sets, which include all possible alternatives that farmers may choose from. The value of modern soybean seed varieties primarily derives from two complementary sources: germplasm (i.e., the underlying genetics accumulated from past generations of selective breeding) and GE traits. The finest possible definition of a product would be in terms of individual varieties. For several reasons, however, analysis at the variety level is not feasible in our context. First, there are simply too many varieties[10]; the implied choice-set by a variety-level product definition would be too large to be estimated with a farm-level mixed logit model. Furthermore, individual varieties have limited geographic presence, as each is bred to be best suited to specific agro-climatic conditions (e.g., latitude). In addition, new varieties are introduced every year, and the life cycle of any given variety is relatively short (four to five years, on average).

It is perhaps more helpful to think of varieties as forming "product lines" over time, as companies introduce improved new varieties that are embedded and built on the genetics of previous varieties. We presume that this continuity is captured by the "brand" (e.g., Asgrow). Varieties marketed by any one brand at different locations may differ, even considerably, but in any one local market one can expect varieties of the same brand to share common characteristics. Hence, we choose to define products by brand, and by whether or not it includes the GT trait.

Specifically, to make the farmers' choice set of the model tractable, but still include as many alternatives as possible, we rely on the 13 distinct brands illustrated in Table 1. Note that we treat the public/university seeds offered by all public sectors as a single brand, named Public.

[10] There are totally 18,420 varieties in our soybean dataset.

These 13 brands account for about 70% of the US soybean seed market, over the period analyzed. All remaining varieties are aggregated into an "Others" group. To account for the important role played by the GT trait, each brand can be associated with two products, depending on whether or not it embed the GT trait. Because Channel only provides GT seeds and Public only offers conventional seeds, we have thus identified a total of 26 distinct products. In any one choice situation, however, a farmer may not have access to all such alternatives. To be more specific about that, we next discuss the definition of "market" used in this study.

3.2 Markets and choice sets

In our model of individual choices, a market is a time-specific location where residing farmers face the same choice set. Following Ciliberto, Moschini, and Perry (2019) we define a market as a CRD-year combination. As noted, CRDs are multi-county sub-state regions identified by the USDA. This market definition is similar to the CRD-level aggregation used in market analyses by some of the major seed companies. Differentiating markets by years is a natural extension, as commercialized varieties evolve over time, and a calendar year contains a natural planting window. In our dataset, we observe a total of 3,791 markets across 233 CRDs. The number of markets in selected years, and the average number of choice alternatives (i.e., products) available to farmers, are provided in Table 2.

Table 2. Markets and Products

Year	Number of Markets	Average number of products Total	GT	Conventional
1996	165	5.60	0.61	4.99
2000	182	9.02	4.95	4.07
2004	174	7.17	5.33	1.83
2008	178	6.06	5.29	0.77
2012	188	5.87	4.88	0.98
2016	189	6.24	4.81	1.42

As noted, a farmer's choice set can contain at most 26 alternatives. The availability of a product in a market is identified by the existence of at least one purchase record. Thus, farmers residing in the same market share the same choice set. Note that, following the introduction of GT varieties, the number of products available to farmers initially increased, but eventually decreased as GT products crowded out conventional products.[11]

3.3 Prices

A common challenge in discrete choice models of individual choices is the construction of prices with transaction data only. The basic problem is that we observe the price for the alternative actually chosen by the individual, but we do not observe the prices of the unchosen alternatives. A typical solution to this problem is to compute average transaction prices and use these prices as the prices that individuals face for each alternative. For example, Goldberg (1995) , in her nested logit model of household automobile vehicles demand, uses the market-level transaction price (net price). On the other hand, Train and Winston (2007) use retail prices in a mixed logit model to study the declining market share of U.S. automakers over time. They argue that, although discounts are common, there seems to be little difference between the discounts offered by American, Japanese, and European manufacturers. In a study of consumer choice behavior in consumer packaged goods markets, where discounts are not common, Keane (1997) uses retail prices, and notes the potential for price endogeneity when using net prices.

In our setting, discounts are common feature of the seed purchasing process—in our dataset, about 63% of 204,697 total observed purchases have a discount. A major reason is the timing of a purchase: farmers who buy earlier are often rewarded with a discount on the listed price.[12]

---

[11] Similar to Train and Winston (2007), the model we develop is a conditional demand model—only soybean seed choices are considered, conditional on the farmer having chosen to plant soybeans on a given plot (i.e., there is no "outside option"). Furthermore, as discussed further beliow, we focus only on new seed purchses (observations where a farmer uses saved seeds are dropped).

[12] The probability of getting a discount is around 30% if a farmer orders their seed in March, April, May, and June, whereas the probability is about 80% if she orders before January. In a logit regression of whether or not the farmer gets a discount, we find that the probability of getting a discount is highest in August and decreases in this order: August, September, October, November, July, December, January, February, March, April, May, June. Additionally, the planting season of soybeans in the United States lasts from June to the end of October (Syngenta 2016).

To account for discounts in the construction of prices for each alternative, we adopt what we term a "contingent" price system. First, in each market, and for any given product, we create two prices: a retail price (the average of all observed retail prices for each product in each market), and a net price (the average of prices after netting out the observed discount in each market). For each farmer, we then identify whether they received a discount. If they did receive a discount, we set the unchosen alternatives' prices to the net prices, and if they did not receive a discount, we set the unchosen alternatives' prices to the retail prices. For the chosen alternative, we assume the farmer faced the price she actually paid, inclusive of the discount, if she received one.

Our goal with this method is to capture the unobserved factors that contributed to the farmer obtaining a discount for the observed choice (for instance, a farmer able to purchase seeds early, and observed to obtain a discount for her seed choice, most likely would have been able to obtain a discount for the other alternatives available to her in the market). We also note that our use of the observed net price for the purchased product (rather than the market price) fits the nature of the problem at hand. Unlike the case of consumer packaged goods, where shelf prices are typically common to all consumers, seed prices are typically negotiated between the farmer and the seller (e.g., dealers or seed companies' representatives).

Finally, we note that the dataset spans 21 years, a long period during which prices changed considerably. Consistent with the homogeneity property of the per-acre profit function, described in what follows, we express all prices in real terms by deflating them by the USDA crop sector index of prices paid. [13]

3.4 Inertia

A major focus of our analysis concerns the possible presence of state dependence in farmers' seed choices. To motivate this perspective, following Dubé, Hitsch, and Rossi (2010), Table 3 reports the purchase rate and repurchase rate of each brand, presented as percentages. The purchase rates are the unconditional probability of choosing each brand, calculated as the market shares over the full time period. Conditional on the previous choice, the repurchase

[13] The Crop Sector Index is published by USDA-NASS Quick Stats. This index takes 2011 as the base year.

rates show the probability of purchasing the same brand again.[14] Note also that the purchase records of a given farmer may not enter the sample in consecutive years. When this occurs, we use the most recent period's records. It is apparent from Table 3 that the repurchase rates are considerably higher than the corresponding purchase rate. In some case, the ratio of the repurchase rate to the purchase rate is extremely high. For example, only 1.28 percent of purchases were for the brand Growmark. Yet, conditional on buying Growmark, an individual had a nearly 76% probability of purchasing it again in the next period. These data are synonymous with persistence or inertia in brand choices over time. Of course, as previously noted, this could because individuals have heterogeneous preferences or because of state dependence.

Table 3. Purchase and repurchase rates of each brand

Brand	Purchase	Repurchase
Asgrow	18.00	78.37
Beck's	2.29	76.51
Channel	1.42	54.88
Croplan	2.78	61.72
DeKalb	3.05	65.37
Golden	3.08	72.81
Growmark	1.28	75.77
Kruger	1.49	74.75
Mycogen	1.78	73.93
NK	7.44	69.86
Other	29.6	85.41
Pioneer	24.12	86.19
Public	0.87	57.05
Stine	2.80	69.74

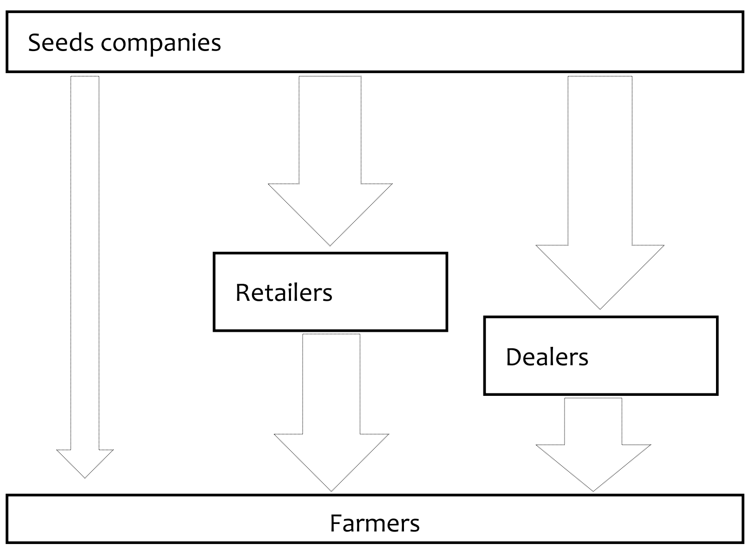
3.5 Purchase sources

The distribution of seed varieties to farmers is highly localized and typically run by independent agents, such as farmer-dealers, farmers' cooperatives, company salespeople, and private wholesalers and retailers (Fernandez-Cornejo 2004). For large farm operations,

[14] In forming Table 3, we do not use the initial purchase records of farmers. Moreover, such initial records are omitted from the regressions in the model, as discussed later in section 5.

seed companies may also sell directly to farmers through their sales representatives (Fernandez-Cornejo, 2004; Syngenta, 2016). The structure of this distribution system is illustrated in Figure 1. Based on this structure, we classify all observed purchases into the three sources that capture the main differences in seed marketing: "sales representative", "dealer", and "retailer". Specifically, a purchase is designated as being from a sales representative if the farmer purchased the seed product directly from the seed company or their representative; it is classified as coming from a dealer if the farmer obtained their seed from a farmer dealer, an independent seed dealer, or if the farmer herself is a dealer; it is classified as coming from a retailer if the products came from any other source, including cooperatives, seed retailers, and grain elevators (seed retailers account for the majority).[15] In our data, 46% of purchases are made from "dealer", 35% are made from "retail", and 19% from "sales representative."

Figure 1. The seed distribution structure in the United States



Note: Adapted from Syngenta (2016, p. 46)

In the econometric model, estimated below, we use the records of purchase sources to generate product-specific marketing variables. These variables potentially control for important sources of unobserved heterogeneity induced by marketing activities.

[15] Retailers differ from dealers in that they typically sell other farm inputs (e.g., fertilizers and pesticides).

4. Model Specification

We develop a seed demand model under the presumption that farmers, on each of their plots, choose the preferred seed alternative to maximize expected profit. The structure of this payoff function depends on the production technology, the prices of output and all other inputs, and, of course, the price of the seed. Ciliberto, Moschini, and Perry (2019) show that, when the production function satisfies two reasonable properties—constant returns to scale in all inputs, and a fixed proportion between land and seed—per-acre expected profit is linear in the (suitably normalized) seed price, i.e.,

$$\pi_{ij} = \pi_{ij}(r, w) - \alpha p_j. \tag{1}$$

Here i indexes the plot to be planted (i.e., the choice situation in what follows), r is the expected price of the output to be produced on this plot, w is the vector of all inputs used in production (except land, meaning that the profit in (1) can be interpreted as returns to land), p_j is the price of seed alternative *j*, and the parameter α captures the (constant) amount of seed per acre (i.e., the seed density).[16]

Given this objective function, the problem for farmer h, in market m, on plot i, can be stated as that of choosing product j such that

$$\max_j \ \pi_{ij}, \ j \in \{1, \ldots, J_m\}, \tag{2}$$

where J_m is the number of available products in the market m pertaining plot i.

4.1 The econometric model

To make this framework operational, we need to parameterize the profit function. In addition to the seed price, which enters linearly, we approximate the other structural determinants (e.g., output and input prices) of the per-acre profit function by a set of seed, market, and

[16] Note that Ciliberto, Moschini, and Perry (2019) express seed prices on a per-acre basis. Here, however, we express seed prices on the typical per-unit measure used in the industry (i.e., per "bag"). Given the assumed fixed proportion between land and seed, the choice of units is immaterial. To translate one into the other note that, in our data, one acre of land on average uses 1.186 units of seed.

farmer-specific variables, in addition to the inertia variable that captures state dependence. Specifically, the per-acre profits for farmer h from product j, for choice situation i in the corresponding market m, are

$$\pi_{ij} = \beta' x_{jm}^{h} + \delta_h' w_{jm} - \alpha p_{jm}^{h} + \gamma_h I_b^h + \nu_{ij} + \varepsilon_{ij}. \tag{3}$$

In this equation, x_{jm}^{h} is a vector of seed characteristics possibly interacted with farmer or market-specific characteristics (the primary seed characteristics include the brand and GT trait), and β is a vector of coefficients that capture the mean impact of each variable in x_{jm}^{h}. The term $\delta_h' w_{jm}$ represents the random coefficients portion of returns: w_{jm} is a vector of seed attributes possibly interacted with market characteristics (e.g., period dummies) and a subset of the variables in x_{jm}^{h}, and δ_h is a vector of normally distributed components with zero mean. Thus, for a variable common to both x_{jm}^{h} and w_{jm}, β represents the mean impact of that variable and δ^h represents the farm-specific deviation from that mean. The variable p_{mj}^{h} is the price of product j in market m for farmer h, and thus the coefficient α represents the impact of price on per-acre returns. Note that to simplify the notation in equation (3), we have not explicitly written that the household (h) and market (m) are uniquely identified by the choice situation (i), and the brand (b) is uniquely identified by the product (j).[17]

Structural state dependence is captured by the indicator variable I_b^h, defined as follows:

$$I_b^h \equiv I\left\{ b \in s\left(b\right)_{t-1}^{h} \right\},$$

where $s\left(b\right)_{t-1}^{h}$ is the set of brands that have been purchased by farmer h in the previous period $t-1$. The indicator variable I_b^h takes value one if the brand associated with alternative j is in set $s\left(b\right)_{t-1}^{h}$. As noted, because the panel is unbalanced, in some cases we do not observe a farmer in consecutive periods. When this occurs, we use the most recent year in

[17] One way to express this explicitly is as $m[i]$, $h[i]$, and $b[j]$.

which the farmer was observed. Finally, ν_{ij} and ε_{ij} are residuals that capture any remaining unobserved variation in profits. We assume that the residual ν_{ij} is normal and correlated with price and that ε_{ij} is i.i.d extreme value. The fact that ν_{ij} is correlated with price is synonymous with the well-known problem of price endogeneity. If this residual is not controlled for in estimation, then the estimated impact of α will be biased.

4.2 Price endogeneity and the control function approach

Price endogeneity is a common issue in the empirical industrial organization literature. At the product-market level, the basic problem is that there are unobserved factors correlated with demand. If firms account for this unobserved characteristic in setting their prices, then the estimated price impacts will be biased. One partial solution to this issue is to include product fixed effects (Nevo 2000), however, there is still the likely possibility that there are product-location-time specific unobserved demand shocks that are correlated with prices. As a result, an estimation procedure that utilizes instrumental variables is usually required.

In our particular setting, there are several potential sources of price endogeneity. As previously noted, we compute prices in a way that accounts for discounts. This may enrich price variation but it could introduce endogeneity bias. More generally, there is the strong possibility of unobserved demand shocks that are correlated with pricing, not only at the product-market level, but also at the farmer level. For example, some farmers may have better relationships with their dealers, which could result in pricing behavior that takes into account a particular farmer's preferences. A third potential source of endogeneity is measurement error in the price variable – the prices we compute for each individual's unchosen alternatives are likely to differ from the prices they actually faced.

In the extant literature, the most common approach to dealing price endogeneity is the "BLP" approach (Train 2009). Most studies that apply this approach, however, use an aggregate discrete choice demand model. The "BLP" approach can still be applied with individual-level data; examples include Berry, Levinsohn, and Pakes (2004), Goolsbee and Petrin (2004), and Train and Winston (2007), but there are two major limitations to using it. First, there are often complications with the contraction part of the algorithm, such as non-convergence (Train 2009). Second, and more importantly, this approach does not control for endogeneity at the farmer level. An alternative approach that resolves both of these issues is the control function approach (Wooldridge 2015; Petrin and Train 2010). Loosely speaking, the control function approach is similar in nature to 2SLS, but can be applied to non-linear models, and is computationally less difficult. Given these benefits, and the fact that we use micro data, we take the control function approach to address price endogeneity. The specifics of this approach are as follows.

Recall that we assume that ν_{ij} is correlated with price and that ε_{ij} is i.i.d extreme value. We assume that prices are determined as follows:

$$p_{mj}^{h} = \gamma z_{ij} + \mu_{ij}, \tag{4}$$

where z_{ijm} includes all variables in equation (3) plus a set of excluded IVs. The residuals μ_{ij} and ν_{ij} are specified as jointly normal. With these assumptions, the per acre profit function can be re-written as:

$$\pi_{ij} = \beta' x_{jm}^{h} + \delta_h' w_{jm} - \alpha p_{jm}^{h} + \gamma_h I_b^h + \lambda_h \mu_{ij} + \varepsilon_{ij}$$

where the distribution of ε_{ij} is still i.i.d extreme value and ν_{ij} has been replaced with $\lambda_h \mu_{ij}$, with the non-correlated component of ν_{ij} asborbed into δ_h, which includes normally distributed brand and trait specific random components. Note also that to allow for additional flexibility in the model, the control function coefficient is specficed as normal: $\lambda_h \sim N(\lambda, \sigma_\lambda)$. In terms of estimation, we first estimate equation (4) and collect the predicted residusals $\hat{\mu}_{ij}$. These residuals are then included as a control variable in the model.

For the IVs, we exploit the fact that soybean seed firms contract out with individual farmers to grow their commercial seed supply for the following year (Lamkey 2004). The terms of contract are set such that the farmer is paid at least what they could have obtained had they planted and sold their own soybeans. This payment will therefore vary in response to changes in expected soybean output prices.[18] A standard proxy for a commodity's expected output price is the futures price corresponding to delivery in the month following the coming season's harvest. Given this, we use the *previous year's* soybean futures price as an instrument for the current year's seed prices. This IV is not only be highly correlated with costs, for the reasons just noted, but will also not affect farmers' relative demand for soybean seed products; i.e., it fulfills the exclusion restriction requirement.[19] To allow for variation across products, we interact futures prices with the brand and GT trait dummies. This is similar to the approach taken by Berto Villas Boas (2007): in her paper on vertical integration in the yogurt market, she creates a set of IVs equal to the interaction of input costs with brand dummies.

4.3 Identification

Before proceeding to the estimation procedure, we informally discuss the intuition of how the model and data separately identify heterogeneity and state dependence. Heterogeneity is entangled with state dependence when pooling data across consumers (Bronnenberg, Dubé, and Moorthy 2019). Consider the following simple example of two farmers over the course of four years. In each year, each farmer makes a single choice between brand A and brand B. An extreme case of AAAA and BBBB as respectively farmer's purchase sequence could be purely the result of heterogeneity or purely the result of state dependence or both.

To disentangle state dependence from heterogeneity, the key is that we need to observe some switching between the brands, specifically caused by choice set variation. Suppose there is a shock in the third period, such as brand B introduces a new variety that significantly

[18] If a seed firm buys its own land and uses that land to grow its commercial supply, the same logic still applies. Fluctuations in expected output prices will change the rental price of land and, therefore, the opportunity cost of seed production.

[19] If the previous year's futures price is correlated with the current year's futures price, it may correlate with a farmers' decision of *which crop* to plant. However, recall that the model we estimate is *conditional* soybean demand model. Thus, there is little reason to think the previous futures price correlates with the demand for particular soybean product. Moreover, we include time specific variables for both brands and GE traits, which should capture any impact of future prices on relative demand.

increases efficiency in production, so the first farmer switches to brand B. If in the fourth period brand A adopts the same variety, state dependence is identified if the first farmer still purchases brand B, whereas the choice consistency should be ascribed to heterogeneity if she switches back to brand A. The case may be too specific in the real world, but it shows the principle: state dependence is identified if the unconditional (of choice set variation) probability of choosing the later brand increases after switch. The extent of state dependence is better measured with rich choice set variation: if the first farmer sticks to brand A with a relative small variation, but switches to brand B with a relative large variation.

Examples of choice set variation used in the prior literature include changes in price, advertising, and the availability of alternatives (Sudhir and Yang 2014). In our context, there are two primary sources of choice set variation. The first source is seed price variation. Relative seed prices flucuate from year to year as different brands try to attract new customers. These fluctations can be in the form of explicit discounts or due to changes in base prices. The second source of variation is changes in product attributes and the availability of alternatives. In particular, the GT trait was not added to all brands at the same time and in the same locations. Moreover, certain brands phased out there conventional varieties faster than others. These changes will have resulted in some farmers either switching to a new brand or trying a new brand. A related source of variation is in the nature of a brand. Seed varieties have relatively short commercial life-cycles. For example, the set of varieties offered under the Asgrow brand in 2000 were quite different from the set of varieties offered in 2010. Thus, for farmers not loyal or partial to Asgrow, there will be ongoing uncertainty about the quality of the brand. From time to time, therefore, such farmers may experiment with a brand like Asgrow to obtain information about it.

4.4 Estimation

The model is estimated using simulated maximum likelihood, as outlined in Hole (2007) and Train (2009). The latent profit function, originally defined in (3), can be written more succinctly as

$$\pi_{ij} = \varphi_{jm}^{h} + \varepsilon_{ij} \tag{5}$$

where φ_{jm}^h includes all components except the IID extreme value error term:

$$\varphi_{jm}^h = \beta' x_{jm}^h + \delta_h' w_{jm} - \alpha p_{jm}^h + \gamma_h I_b^h + \lambda_h \mu_{ij} \tag{6}$$

For a *given* realization of φ_{jm}^h, the probability that farmer h chooses alternative j in choice situation i is given by the familiar logit expression

$$L_{ij}\left(\theta^h\right) = \frac{\exp\left(\varphi_{jm}^h\right)}{\sum_{n=1}^{J} \exp\left(\varphi_{nm}^h\right)} \tag{7}$$

where $\theta^h = \{\beta, \delta_h, \alpha, \gamma_h, \lambda_h\}$ is the vector of coefficients to be conditioned on. For each farmer, we observe a *sequence* of choices. The probability of that sequence for farmer h, conditional on θ^h, is given by the product of the logits:

$$L^h\left(\theta^h\right) = \prod_{i=1}^{I^h} L_{ij}\left(\theta^h\right) \tag{8}$$

where I^h indexes farmer h's final choice, and the set $\left\{1, 2, \ldots, I^h\right\}$ represents farmer h's choice sequence (on average, a farmer makes 23.6 purchases overall and 4 purchases per year).

To obtain the unconditional probability of an individual's choice sequence we need to integrate out the random components, denoted by R, in θ^h — $\delta_h, \gamma_h, \lambda_h$. As outlined above, we assume that each of these random coefficients follows an independent normal distribution whose mean and variance are estimated (the means of δ_h are zero as they are already captured by β, as noted). Thus, the unconditional probability of farmer h's sequence of observed choices can be written as

$$L^h = \int L^h(R)\phi(R|M, SD)\,dR, \tag{9}$$

where M is the mean vector of random coefficients R and SD is the corresponding standard deviation. The probability is simulated using 200 Halton draws for any given value of M and SD.[20] Specifically,

$$\tilde{L}^h = \frac{1}{D}\sum_{d=1}^{D} L^h\left(R^d\right) \tag{10}$$

where D is the number of draws and is equal to 200 (d indexes each draw). R^d is a realization of the random coefficients from the dth draw from a given normal distribution. As noted by Train (2009), equation (10) is an unbiased estimator of L^h by construction. The log-likelihood for the model is

$$\ln L = \sum\nolimits_{h=1}^{H} \ln \tilde{L}^h \tag{11}$$

The parameters are estimated by maximizing the log-likelihood in STATA using 200 Halton draws. Specifically, we are utilize the user written "mixlogit" package by Hole (2007), outlined in Cameron and Trivedi (2005).

5. Results

In this section, we first discuss our preparation of the estimation data, as well as present some basic summary statistics. We then present estimation results for the basic conditional logit model and the full mixed logit model. In section 5.2, using the coefficient estimates from the full mixed logit model, we compute and discuss WTP distributions for the main variables of interest. We then compute mean own-price and cross-price demand elasticities. Finally, in section 5.4, we consider the robustness of the state dependence result by conducting a reshuffling procedure along the lines of Dubé, Hitch, and Rossi (2010).

[20] Halton draws are used to approximate the distribution of random coefficients. Because there is no closed form expression for equation (9), we simulate the equation by taking the mean of Halton draws. As noted in Train & Winston (2007), Train (2009), Petrin & Train (2010), 100 Halton draws is more efficient than 1000 random draws and 200 Halton draws is sufficient for simulation.

5.1 Data Preparation and Summary Statistics

In total, the dataset contains 213,062 seed purchase records for 28,017 farmers. We clean and reformat the dataset for estimation of the conditional and mixed logit regressions following the steps listed in Table 4.

Table 4. Data cleaning and reformation

Deletion	No.
Original soybean choices	**213,062**
Purchase records if purchase source is "From my own farm", "I'm a seed grower", or "New seed that was left over from last year"	6,890
Purchase records if the seed is not newly purchased	1,367
Purchase records if the product is "Public" with GT trait	108
Choices in analysis	**204,697**
(Following deletions will not affect purchase history)	
Purchase records of zero net prices	831
Purchase records of year 1996, 1997, 1998	21,924
Purchase records in markets with only one alternative	584
(Following deletions will not affect available alternatives in one market)	
Purchase records of all farmers' first-time recorded purchase	62,548
Purchase records of all farmers if they show up in less than or equal to 3 years in the total time span	28,546
Choices in regression	**90,264**
(Expand each choice by available products in the local market)	
Observations in regression	**1,057,637**

We drop all cases in which a farmer did not purchase a new soybean variety. These cases include the following purchase classifications: "From my own farm", "I'm a seed grower", or "New seed that was left over from last year." Next, because public varieties have note included the GT trait, we assume the small number of cases in which they were associated with the GT trait was an error, and therefore drop these observations (108 in total). This leaves 204,697 purchase records, which are termed "Choices in analysis" in Table 4. We highlight this number and this step because this is the dataset that provides all information used in the empirical analysis of this paper. However, the first three years of data (1996-98) are exclusively used to build farmers' purchase history. In addition, the first year a farmer appears in the data (for many farmers this is after 1998) is used to create the "state dependence" variable, and thus such observations are not use in the logit regressions. Again, however, these records still enter the model through certain explanatory variables, such as the initial brand choices, the state dependence terms, and the marketing variables. Moreover, for robustness, we drop purchase records of farmers who appear in the sample only three or less year. The end result of this process yields 90,264 purchase records that provide the estimation data for the logit regressions (termed "Choices in regression" in Table 4). For the model to be estimated, the data needs to be further expanded such that, for each individual choice, there is also a row of information for each unchosen alternative in the corresponding market; this expansion results in 1,057,637 "observations". Descriptive statistics, calculated from the dataset of "Choices in analysis", are provided in the previous Tables 2 and 3, whereas Table 5 describes the dataset of "Choices in regression."

Table 5 shows that a farmer generally chooses about two brands (and also two products) per year, and the number of chosen brands doubles for the whole observed period of a farmer, suggesting some brand switching, which aids identification of structural state dependence (further discussed in section 4.2). Table 5 also demonstrates that, on average, a farmer chooses among 11.7 alternatives, with a minimum of two (note that we drop records in markets with only one alternative) and a maximum of 23 (recall that our product definition results in 26 possible alternatives). For a typical farmer, we observe about 8 years of data (recall the minimum is four, as discussed in the foregoing, and the maximum is 21, i.e., a farmer appears in the sample in every year over the 1996-2016 period).

Table 5. Descriptive information

Variable	Mean	SD	Min	Max
No. of alternatives in one market	11.717	4.393	2	23
No. of chosen brands of a farmer	3.730	1.831	1	11
No. of chosen brands in one year's purchase of a farmer	1.901	1.009	1	8
No. of chosen products of a farmer	4.659	2.538	1	17
No. of chosen products in one year's purchase of a farmer	1.994	1.084	1	8
No. of recorded years of a farmer	7.984	3.689	4	21
Retail price	44.924	10.900	8.757	89.646
Net price	40.734	9.050	8.757	85.167
Discount	4.189	4.321	0	37.363

Note: the mean is calculated by averaging over all purchase records, rather than averaging over the markets or farmers

5.2 Model Variables

Recall that the regression model includes two primary sets of explanatory variables, represented by x_{jm}^{h} and w_{jm}. Within the vector x_{jm}^{h}, there are three types of variables: (i) a set of brand and trait intercepts, possibly interact with each and time; (ii) initial condition variables, and (iii) a set of marketing variables and their interactions with individual-specific purchasing experiences. Each of these three types of variables constitute 52, 15, and 6 variables, respectively. Further details are as follows.

In the first set of variables in x_{jm}^{h}, the brand and trait intercepts capture the average profit gains of each brand and GT trait. In the estimation, we classify our regression timespan 1999-2016 into three periods 1999-2004, 2005-2010, 2011-2016 with the same length and interact the three periods with brand and trait intercepts to account for any time-variant effect. In the model they are coded as GT 1999-2004, GT 2005-2010, GT 2011-2016, and “brand”, “brand” 1999-2004, “brand” 2005-2010 (“brand” refers to the brands listed in Table 1, except Public, the baseline brand; we further take "brand" 2011-2016 as the baseline, which appears in the model directly as "brand" and we only report them in the following Table 6). For time-variant brand effects, they can be caused by introduction of different new varieties over time, or any brand-specific changes that happen nationwide. The time-variant trait effect may come from the commercialization of new GT patent, the emergence of glyphosate-tolerant weed, etc. Similarly, we only capture the national effect rather than the CRD-specific effect. In the conditional logit environment, case-specific variables, like time or CRD region, will not affect the choice decision if they influence all alternatives in the same way, so they come in the model by interacting with alternative-specific brand or trait intercepts. In our case, time-variant brand effects are explicitly model, whereas CRD-variant brand or trait effects go to the error terms. Finally, the interaction variables of brand and trait, coded as "brand"_GT in the model, capture the possible different GT propensity over brands, which can materialize if one brand have more varieties associated with GT trait or bundle the GT trait with other desirable seed attributes, like soybean cyst nematode resistance. There are 11 interactions of brand and GT trait, excluding Public_GT, Channel_GT, and Kruger_GT for collinearity issues.[21]

The second set of variables in x_{jm}^{h} capture the initial conditions of a farmers' choice sequence. Following the idea of Wooldridge (2005), we include initial states as extra explanatory variables to account for the initial conditions problem of correlation between unobserved heterogeneity and the initial state. By further assuming the unobserved heterogeneity is normally distributed conditional on the initial states, we can integrate the conditional heterogeneity out by simulation and get what Wooldridge called "conditional" maximum likelihood estimates. Specifically, we add 14 brand-specific initial brand choices variables, as "brand"_ini in the model, and one brand propensity variable, iniacre in the model, constructed as the percentage of land planted with a brand in the initial state. These 15 variables are both

[21] All Public seeds are conventional and all Channel seeds are associated with GT trait, so Public_GT and Channel_GT are collinear with the corresponding brand intercepts. We further take Kruger_GT as the baseline. If not, this set of variables will be collinear with the trait intercepts.

alternative-specific, coding only for brands purchased in the first period, and farmer-specific, different across farmers but same for all choices made by the same farmer.

Finally, the last set of variables in x_{jm}^{h} is to control for possible heterogeneity brought by marketing activities. Corresponding to the three purchase sources, we construct three marketing dummy variables—dealer, rep, and retailer in the model. For any given market, we say a source is active for a product (and thus the corresponding marketing variable take value 1) if the corresponding brand is recorded to be purchased from this source in this market in any one of last three years. Thus, these marketing variables are market-product-specific, so they are same for all alternatives of a product among the market.[22] Note that we do not use current year's records to construct these marketing variables to avoid potential endogeneity issues (the purchase source and the choice decision are made simultaneously). We further interact these three variables with three individual-specific variables, coded as ind_dealer, ind_rep, ind_retailer, to capture farmers' heterogeneous responses to the marketing variables. For each source, the corresponding individual-specific variable tells whether she has experience the purchase source in any previous period.

To control for unobserved heterogeneity, our model allows heterogeneous responses to variables in w_{jm}. However, to keep the computational burden of the model within tractable boundaries, w_{jm} does not contain all variables in x_{jm}^{h}. Specifically, w_{jm} consists of the GT trait intercepts over time, GT 1999-2004, GT 2005-2010, GT 2011-2016 and the brand intercepts, "brand", a total of 16 variables.

Additionally, the cost instruments are constructed as interactions of futures prices with brand and trait dummies. As noted by Fernandez-Cornejo and Spielman (2002), the cost of seed production can be contracted as adjusted yields times the futures prices of the product for contract farmer growers. Specifically, the futures prices are constructed from the futures

[22] Note here the marketing variables are market-product-specific rather than farmer-product-specific. To generate farmer-product-specific marketing variables, the purchase source information is needed even for not chosen alternatives, which is not available in our case.

contract with a delivery month of November (right after the harvest season) as the average daily closing prices from January to March (before the planting season) as in Kim and Moschini (2018), who use futures prices as expected prices for farmer producers. The price records are from Quandl.[23] We also deflate the futures prices by the USDA crop sector index of prices paid, consistent with all other prices in the model.

5.3 Estimation Results

Table 6 presents estimation results for four specifications: the first three columns contain results for three different cases of the basic conditional logit model (i.e., the logit model without random coefficients), and the fourth and fifth columns contain the results for the full mixed logit model. For the conditional logit results, the first column ("No IV") provides results for the case when price is assumed to be exogenous, the second column ("Main") provides results with IVs included (i.e., the control function approach), and the third column "No state" omits the state dependence variable.

Table 6. Estimation Results

	Conditional logit			Mixed Logit	
	No IV	Main	No state	Main, coef	Main, SD
	b/(se)	b/(se)	b/(se)	b/(se)	
price	-0.064***	-0.191***	-0.175***	-0.305***	
	(0.003)	(0.015)	(0.014)	(0.025)	
state	2.344***	2.316***		2.065***	0.932***
	(0.039)	(0.040)		(0.045)	(0.035)
GT 1999-2004	1.524***	3.141***	2.942***	5.217***	2.153***
	(0.137)	(0.221)	(0.207)	(0.388)	(0.099)
GT 2005-2010	2.881***	4.573***	4.355***	8.459***	3.260***
	(0.139)	(0.252)	(0.235)	(0.507)	(0.206)
GT 2011-2016	2.122***	3.224***	3.076***	6.989***	3.725***
	(0.157)	(0.186)	(0.181)	(0.371)	(0.265)
Asgrow	(0.262)	2.181***	2.493***	5.133***	0.588***
	(0.202)	(0.392)	(0.421)	(0.632)	(0.055)
Beck's	0.848**	3.813***	3.758***	5.730***	1.990***
	(0.269)	(0.449)	(0.476)	(0.752)	(0.142)
Channel	0.230	2.669***	2.631***	2.588***	2.896***
	(0.225)	(0.388)	(0.416)	(0.693)	(0.300)
Croplan	0.274	3.076***	3.009***	5.326***	-1.002***
	(0.206)	(0.405)	(0.430)	(0.650)	(0.091)
DeKalb	-1.452***	0.450	0.041	2.440***	0.275*
	(0.237)	(0.369)	(0.397)	(0.559)	(0.131)

[23] See https://www.quandl.com.

Golden	-1.447***	0.641	0.024	2.448***	0.627***
	(0.312)	(0.427)	(0.463)	(0.607)	(0.110)
Growmark	-0.388	2.224***	1.947***	4.688***	0.863***
	(0.249)	(0.420)	(0.445)	(0.661)	(0.159)
Kruger	(0.471)	1.930***	1.576***	3.392***	1.542***
	(0.272)	(0.442)	(0.465)	(0.917)	(0.315)
Mycogen	-0.148	2.221***	1.891***	4.111***	1.420***
	(0.229)	(0.368)	(0.409)	(0.609)	(0.144)
NK	0.131	2.550***	2.520***	5.022***	0.725***
	(0.182)	(0.361)	(0.389)	(0.568)	(0.070)
Other	0.268	2.491***	2.718***	4.909***	0.693***
	(0.170)	(0.341)	(0.369)	(0.547)	(0.045)
Pioneer	0.050	2.484***	2.658***	5.318***	-0.356***
	(0.178)	(0.359)	(0.384)	(0.592)	(0.091)
Stine	0.093	2.391***	2.018***	4.650***	0.528*
	(0.238)	(0.396)	(0.442)	(0.603)	(0.223)
dealer	-0.159**	-0.151**	-0.101	-0.073	
	(0.054)	(0.054)	(0.061)	(0.058)	
ind_dealer	0.256***	0.237***	0.450***	0.179*	
	(0.063)	(0.063)	(0.070)	(0.070)	
rep	0.050*	-0.019	0.019	-0.030	
	(0.025)	(0.025)	(0.030)	(0.030)	
ind_rep	0.088*	0.112**	0.355***	0.119**	
	(0.040)	(0.040)	(0.052)	(0.044)	
retailer	0.053	0.009	0.024	0.029	
	(0.049)	(0.050)	(0.059)	(0.061)	
ind_retailer	-0.027	0.012	0.251***	0.069	
	(0.057)	(0.058)	(0.067)	(0.072)	
control		0.127***	0.106***	0.207***	0.184***
		(0.014)	(0.013)	(0.025)	(0.006)
LL	-124090.968	-123996.633	-151032.646	-108463.555	
N	1057637	1057637	1057637	1057637	

Note: standard errors are clustered at CRD level. *p<0.05, **p<0.01, ***p<0.001. The sign of the estimated standard deviations is irrelevant: consider them as positive.

Looking across the estimation results, the coefficients are generally estimated with high precision, especially when the control function is used. The importance of the control function approach is also demonstrated by the difference in results between the "no IV" and the "Main" estimation results. The price coefficient decreases substantially from about -0.06 to -0.19, suggesting that endogeneity was indeed present. Many of the brand coefficients become statistically significant and positive in magnitude. Given that public varieties are the reference product; this is more consistent with expectations. The positive and significant estimate for the control residual ("control") also suggests the existence of unobservable factors positively correlated with demand, a finding similar to that found in Petrin and Train (2009). To demonstrate the influence of state dependence on the other coefficients, the third column "No state" drops the state dependence term from the model. Here we see a sharp decrease in the log-likelihood, on the order of 25%, and there is some increase in the marketing variables. The remaining variables are relatively unchanged.

As noted, the final two columns contain estimation results for the full mixed logit model, our preferred specification and the specification we later use to conduct some counterfactual exercises. The "Main, coef" and "Main, SD" reports the means and standard deviations of the corresponding coefficient. The full model results demonstrate the importance of specifying random coefficients for state dependence, the GT dummies, and the various brand dummies. As evidence of this, the log likelihood further decreases by more than 10% and all variance parameters are significant and large in magnitude. Taste variation is particularly large for the GT traits and the brands Channel, Kruger, and Mycogen. The mean estimate for state dependence also decreases, suggesting some previous bias by not allowing for unobserved heterogeneity. Nonetheless, it still remains quite large in magnitude.[24] This could be because it is *truly* large in magnitude or because our specification of unobserved heterogeneity is still not rich enough to capture the observed purchase patterns. We consider this issue further below.

The mean coefficients for the GT and brand variables generally conform to expectations. All suggest a positive impact on the return to a soybean product. Further discussion of these variables and their economic importance are discussed in the next section where compute willingness-to-pay (WTP) distributions. The coefficients for the marketing variables are generally smaller in magnitude and less statistically significant. They do suggest that an active dealer or sales representative in the market can increase farmers' probability of choosing the brand only if the farmer has purchased from the respective source before. If the farmer has no relevant source experience, an active dealer or sales representative of the brand in the local market may even decrease his probability of choosing the brand, although not

[24] The coefficient of 2.065 is equivalent to an average marginal effect of 9 percentage points. That is, the average probability of choosing a particular brand will increase by 0.09 if it was purchased in the most recent observed period.

significantly. We also show that an active retailer source has little effect on the farmer's probability of choosing the brand, regardless of her previous experience.

Table 6 does not contain all of the estimated coefficients; the full set of results are provided in the Appendix. Among the omitted coefficients, there are a few noteworthy results. First, essentially all initial conditions—the brand-specific dummy variables coding for whether a farmer purchased the brand in their first observed year—were positive and significant (except for Asgrow and Other), with the effects being the largest for smaller brands such as Beck's, Growmark, Kruger, and Mycogen, and Channel. Moreover, in including the initial conditions we found that the state dependence coefficient decreased significantly, suggesting these variables capture an important source of unobserved taste variation not captured by the random coefficients. A second set of results omitted from the table are the brand-GT trait interaction coefficients. Here we found that the GT trait is more valued under some brands than others. For example, the GT trait is most valued when sold under the Asgrow brand and least valued (relative to conventional varieites) in Beck's varieties. Finally, the brand-period interaction coefficients, also omitted, indicate that, compared to the final period, all brands were valued less (relative to public varieties) in the first two periods. This is consistent with the decline in demand for public varieties.

5.2 WTP distributions

The estimated coefficients presented in Table 6, per se, are not terribly informative about the economic importance of the various factors that impact the profitability of soybean varieties. Therefore, we use the coefficient estimates from the mixed logit model to report the WTP distributions for the main variables of interest: the structural state dependence term, the GT coefficient, and the brand coefficients.

The WTP for an attribute measures the maximum amount ($/unit) that a farmer is willing to part with for that characteristic. The WTP distribution for each attribute is obtained by dividing its mean and variance coefficients by the price coefficient (Train 2009).[25] Because the random coefficients are normally distributed, and the price coefficient is a constant, the WTPs for each attribute are also normally distributed. The mean and variance of the WTP distributions for each variable of interest are reported in Table 7.

Table 7. Estimated WTP distributions

WTP	Mean	SD
state	6.765	3.053
GT 1999-2004	17.089	7.053
GT 2005-2010	27.709	10.679
GT 2011-2016	22.892	12.202
Asgrow	16.814	1.926
Beck's	18.770	6.519
Channel	8.477	9.486
Croplan	17.447	3.282
DeKalb	7.993	0.901
Golden	8.020	2.054
Growmark	15.357	2.827
Kruger	11.110	5.051
Mycogen	13.467	4.651
NK	16.452	2.375
Other	16.081	2.270
Pioneer	17.422	1.166
Stine	15.231	1.730

We begin with the value of structural state dependence. On average, a farmer is willing to pay about 6.77 ($/unit) more for a brand if they purchased it in the previous period. Given that the mean price for a unit of soybeans is $45/unit, the state dependence effect is about 15% of the average, a relatively large effect and quite close to what Dubé, Hitsch, and Rossi (2010) find in the consumer-packaged goods industries.[26] Put differently, experience is important: for two

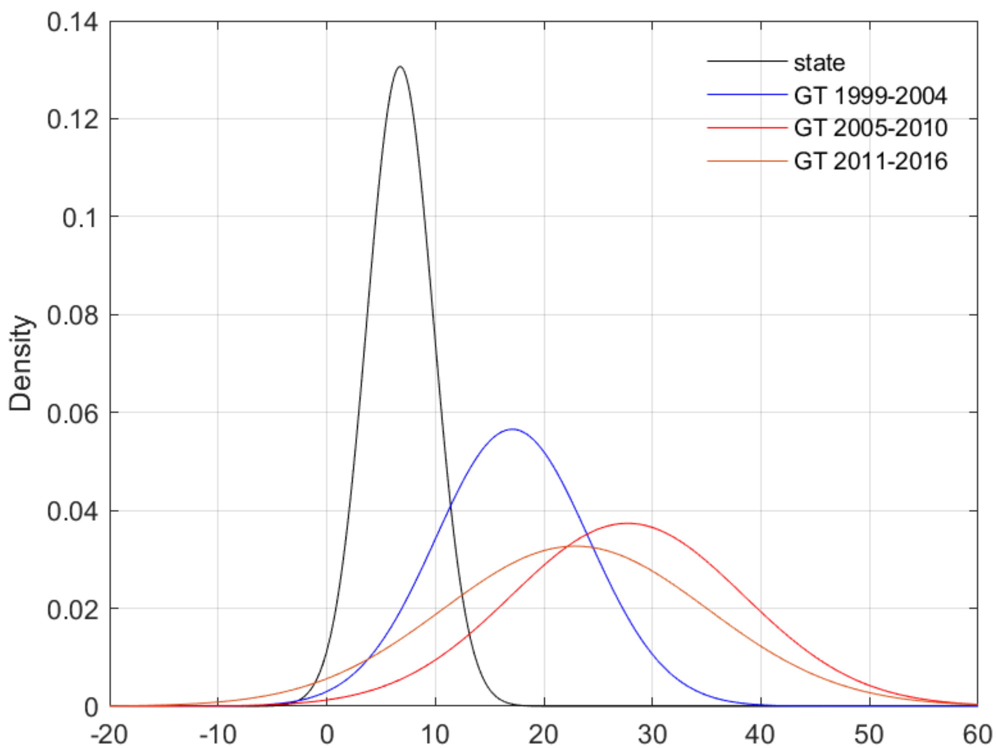
[25] We take advantage of the constant price coefficient. If the price coefficient is random, recovering the WTP distribution is trickier. See more discussions in Train & Weeks (2005) and Louviere et al. (2005).

[26] As shown in Table 5, the average retail seed price is $45/unit—$37/unit over the period 1999-2004, $44/unit in the period 2005-2010, and $53/unit in the period 2011-2016. All prices are deflated by the crop sector index of prices paid. To give these number additional context, the U.S. average soybean yield in 2018 was 51.6 bushels per acre and soybean price was $9.15/bushel, implying an average total gross revenue $398/unit (measured in 2018 dollar, adjusted by the unit/acre ratio).

otherwise equally valued products, a farmer is willing to pay significantly more for the one he has experience with. There is also significant heterogeneity in state dependence. A substantial number of farmers value an additional past experience at more than $9, whereas some farmers obtain little value from an additional experience.

To give a visual depiction of the WTP distributions, Figure 2 provides drawn normal distributions for the state dependence and GT trait coefficients. This demonstrates that the value of structural state dependence is positive for nearly all farmers. The same can be said for the value of GT traits, and the vast majority of farmers place positive value on GT during all sub-periods. We can also see that the value of the GT trait is larger on average and more dispersed than the value of state dependence. This suggests that some farmers place very high value on the GT trait whereas other farmers do not. Consider, for example, that in the final period about 15% of farmers valued the GT trait at $35 or more, whereas another 15% valued the GT trait at $10 or less.

Figure 2. WTP distributions for state and trait



The three different WTP distributions also demonstrate interesting changes over time. In the first period, 1999-2004, the average WTP is about 17.09 ($/unit), this increases to 27.71 ($/unit) in the second period (2005-2010), and then decreases to 22.89 ($/unit) in the third period (2011-2016) (in $/acre terms, these values are $20.27/acre, $32.86/acre, and $27.15/acre, respectively).[27] These changes essentially reflect the changing rate of GT soybean adoption over time. As noted in CIliberto, Moschini, & Perry (2019), the observed increase after the period 1999-2004 was likely the result of falling glyphosate prices (a complementary input to GT soybeans), rising output prices, learning, and/or an increasing number of varieties with GT. However, in the final period, the mean WTP for GT decreases and its variance increases (the orange line). This is consistent with some recent developments in the efficacy of GT soybeans. In recent years, glyphosate weed resistance has become increasing problematic (Perry, Ciliberto, Hennessy, and Moschini, 2016). Some farmers have responded by switching to non-GT varieties or by increasing glyphosate application rates (Perry, Hennessy, Moschini, 2019). Glyphosate weed resistance also varies considerably across the U.S., which may explain the increasing variance in farmers' WTP.

The remaining entries in Table 7 represent farmers' mean WTP for the each of the brands *relative* to public varieties in the last period (2011-2016). As expected, farmers are willing to pay a signficant premium for a branded product, with Asgrow, Beck's, and Pioneer having some of the highest WTPs. There is also considerable heterogeneity. Generally, the larger brands such as Asgrow, Pioneer, and NK have the most concentrated WTPs, whereas some of the mid-size to smaller brands like Beck's and Channel have widely distributed tastes.

5.3 Demand elasticities

A major advatange of our framework, as contrasted with the basic logit model, is that it can capture rich substitution patterns between seed varieties. To demonstrate these patterns, we compute and report simulated mean own-price and cross-price demand elasticities for each product. To compute these elasticities, we compute predicted market shares for each product

[27] In Ciliberto, Moschini, and Perry (2019), the WTP for the GT trait in soybeans was $16.68/acre in 1996-2000, $23.25/acre in 2001-2006, and $24.66/acre in 2007-2011.

j, denoted S_j. To obtain these shares, we first predict farm-level probabilities for each product using the estimated coefficients from the full mixed logit model. Because this model includes random coefficients, this is done through simulation. Given the individual level probabilities, we then aggregate to the product-market level using the observed number of purchased units for each farmer as weights. To generate the elasticities, we change the price for each product k by a small amount $p'_k - p_k$, and then recompute the aggregate prediced market share for product j, denoted by S'_j. The elasticity of demand for product j with respect to a change in the price of product k is given by:

$$e_{jk} = \frac{\Delta\left(S'_j - S_j\right)}{\Delta\left(p'_k - p_k\right)} \frac{p_k}{S_j}, \tag{12}$$

where e_{jj} is the own-price elasticity of demand and $e_{jk}\left(j \neq k\right)$ represents the cross-price elasticity of demand j for product k.

Table 8 contains the full matrix of elasticities. Each product j is listed in the first column and each product k is listed in the top row. Thus, the own-price and cross-price elasticities of demand for product j with respect to a change in the price of product k are reported in the corresponding row. For example, the top-left entry of "-6.20" corresponding to Asgrow0 and AS0 is the own-price elasticity of demand for Asgrow0; the value to the right of this, "0.03", corresponding to Asgrow0 and BE0, is the cross-price elasticity of demand for Asgrow0 with respect to a change in the price of Beck's0. Note that we use "0" to represent conventional products and "1" to represent GT products. In the first row of Table 8, to save space, we only use the first two letters of each brand. We further divide Table 8 into 4 sub-matrices—the top left panel contains elasticities for conventional products with respect to other conventional products, the top right panel contains elasticities for conventional products with respect to GT products, the bottom left panel contains elasticities for GT products with respect to conventional products, and the bottom right panel contains elasticities for GT products with respect to GT products. Grouping it this way allows us to see some clear patterns in the elasticities.

The orange cells in Table 8 contain the own-price demand elasticities for each product. They are all negative and highly elastic, typically ranging from about -5 to as high as -8.8. The values are also similar to the estimated mean own-price elasticties reported in CIliberto, Moschini, & Perry (2019). In their preferred model, they find a mean own-price elasticity of -7.04 for corn and soybean products. We also note that GT products are slightly more elastic than conventional products, which may in part be the result of higher prices for GT products.

The blue cells highlight cross-price elasticities for different products marketed with the same brand. Because not all brands posses both GT and conventional products (Channel and Public), there is some asymmetry along the alignment of the blue cells. Cells highlighted in green identify the closest substitute for each product k: this is simply the cell with the highest value in each column (excluding the own-price elasticities). If this cell *also* happens to be the product with the same brand, then it is highlighted in blue-green.

Three intuitive regularities emerge from the cross-price elasticies. First, generally speaking, a farmer is more likely to substitute between products that contain the same trait. We term this the "trait effect". Put differently, if a product with GT is a farmer's most preferred variety, it is significantly more likely that their next preferred product also has GT. This can be seen by the fact that the upper left and lower right blocks of cross-price elasticities are typically larger compared to the lower left and upper right blocks (the main exception to this is the cross-substition from GT to conventional products of the same brand). Second, individuals are typically more likely to substitute products of the same brand. We call this the "brand effect". Consider, for example, how individuals substitute from ASo to products with GT (the first column in the lower left block of Table 8). Among all such products, farmers are most likely to substitue to Asgrow1: the value of 0.07 exceeds all other values in the lower left panel. Notice, however, that the cross-price elasticites for all conventional products from ASo are greater than 0.07. Thus, in this case, the trait effect dominates the brand effect. More generally, for conventional products, the trait effect usually dominates the brand effect (though not always). This is evidenced by the fact that the majority of green cells for conventional products are in the upper right block. Conversely, for GT products, the closest substitute is almost always the identically branded conventional version. For example, the closest

Table 8. Own-price and cross-price elasticities of demand

elasticity	AS0	BE0	CR0	DE0	GO0	GR0	KR0	MY0	NK0	OT0	PI0	PU0	ST0	AS1	BE1	CH1	CR1	DE1	GO1	GR1	KR1	MY1	NK1	OT1	PI1	ST1
Asgrow0	-6.20	0.03	0.06	0.19	0.21	0.06	0.05	0.07	0.12	0.77	0.52	0.16	0.16	2.50	0.04	0.01	0.08	0.26	0.20	0.06	0.06	0.05	0.17	0.93	0.72	0.13
Beck's0	0.14	-6.94	0.05	0.01	0.02	0.05	0.00	0.01	0.00	0.75	0.69	0.04	0.16	0.84	1.92	0.12	0.09	0.04	0.03	0.07	0.01	0.05	0.27	0.86	0.93	0.09
Croplan0	0.16	0.04	-6.14	0.14	0.15	0.01	0.06	0.11	0.13	0.84	0.55	0.16	0.22	0.75	0.04	0.03	0.99	0.23	0.19	0.02	0.07	0.08	0.25	1.01	0.87	0.17
DeKalb0	0.32	0.04	0.10	-6.22	0.23	0.06	0.06	0.10	0.12	0.78	0.56	0.17	0.22	0.70	0.04	0.00	0.08	1.58	0.24	0.06	0.06	0.05	0.16	0.94	0.74	0.14
Golden0	0.25	0.03	0.07	0.16	-5.69	0.05	0.05	0.08	0.12	0.76	0.46	0.13	0.19	0.67	0.04	0.00	0.09	0.26	1.59	0.06	0.07	0.06	0.17	1.02	0.68	0.15
Growmark0	0.29	0.03	0.02	0.14	0.15	-5.79	0.05	0.07	0.08	0.61	0.52	0.11	0.16	0.64	0.05	0.03	0.05	0.20	0.19	1.85	0.06	0.04	0.13	0.76	0.74	0.12
Kruger0	0.23	0.00	0.11	0.14	0.20	0.05	-4.97	0.09	0.12	0.60	0.42	0.08	0.20	0.55	0.01	0.00	0.10	0.23	0.21	0.06	1.79	0.06	0.16	0.76	0.59	0.14
Mycogen0	0.19	0.01	0.11	0.15	0.18	0.05	0.06	-5.29	0.13	0.56	0.42	0.14	0.17	0.60	0.02	0.00	0.08	0.24	0.22	0.06	0.06	1.04	0.15	0.96	0.68	0.14
NK0	0.24	0.01	0.10	0.13	0.19	0.04	0.05	0.10	-5.91	0.79	0.51	0.17	0.13	0.63	0.02	0.00	0.12	0.25	0.25	0.05	0.06	0.06	1.49	1.01	0.76	0.13
Other0	0.13	0.05	0.06	0.06	0.09	0.03	0.02	0.03	0.06	-4.74	0.38	0.13	0.13	0.72	0.05	0.05	0.11	0.14	0.13	0.04	0.05	0.06	0.26	2.76	0.73	0.11
Pioneer0	0.18	0.07	0.07	0.09	0.12	0.04	0.03	0.05	0.09	0.71	-5.65	0.12	0.14	0.70	0.06	0.04	0.10	0.18	0.15	0.06	0.05	0.05	0.24	0.86	2.62	0.11
Public0	0.19	0.02	0.08	0.11	0.12	0.03	0.02	0.07	0.10	0.88	0.43	-3.40	0.15	0.87	0.03	0.02	0.14	0.22	0.19	0.06	0.04	0.05	0.31	1.41	0.89	0.14
Stine0	0.19	0.07	0.10	0.14	0.17	0.05	0.05	0.07	0.08	0.94	0.52	0.15	-5.97	0.83	0.04	0.05	0.11	0.21	0.22	0.07	0.08	0.08	0.25	1.01	0.86	1.03
Asgrow1	0.07	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.01	0.14	0.08	0.02	0.02	-6.76	0.12	0.14	0.28	0.25	0.16	0.09	0.08	0.12	0.66	1.93	1.87	0.19
Beck's1	0.02	0.32	0.01	0.00	0.01	0.01	0.00	0.00	0.00	0.10	0.08	0.01	0.01	1.31	-6.67	0.16	0.19	0.14	0.08	0.10	0.03	0.08	0.49	1.64	1.56	0.16
Channel1	0.01	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.07	0.01	0.02	1.56	0.16	-7.29	0.24	0.07	0.02	0.09	0.08	0.17	0.69	1.71	1.80	0.19
Croplan1	0.02	0.01	0.08	0.01	0.01	0.00	0.00	0.00	0.01	0.15	0.07	0.03	0.02	1.95	0.13	0.15	-8.79	0.25	0.20	0.05	0.11	0.16	0.88	2.21	1.95	0.27
DeKalb1	0.05	0.01	0.02	0.14	0.04	0.01	0.01	0.02	0.02	0.18	0.12	0.04	0.03	1.86	0.11	0.03	0.25	-8.24	0.41	0.10	0.13	0.11	0.58	2.01	1.61	0.26
Golden1	0.04	0.01	0.02	0.03	0.26	0.01	0.01	0.02	0.03	0.18	0.11	0.03	0.04	1.24	0.08	0.00	0.21	0.44	-7.34	0.10	0.13	0.10	0.51	1.84	1.33	0.25
Growmark1	0.05	0.02	0.01	0.01	0.02	0.25	0.01	0.01	0.02	0.14	0.11	0.03	0.03	1.59	0.15	0.14	0.14	0.26	0.23	-8.03	0.14	0.09	0.56	1.61	1.71	0.26
Kruger1	0.03	0.00	0.01	0.01	0.03	0.01	0.16	0.01	0.02	0.14	0.08	0.02	0.03	1.26	0.04	0.10	0.24	0.30	0.29	0.12	-6.77	0.13	0.55	1.55	1.40	0.26
Mycogen1	0.03	0.01	0.02	0.01	0.01	0.00	0.00	0.11	0.01	0.16	0.08	0.02	0.03	1.64	0.09	0.20	0.30	0.18	0.17	0.07	0.10	-7.85	0.71	1.90	1.77	0.24
NK1	0.02	0.01	0.01	0.00	0.01	0.00	0.00	0.00	0.05	0.14	0.07	0.03	0.02	1.72	0.11	0.16	0.32	0.21	0.17	0.08	0.09	0.14	-7.66	2.06	1.92	0.22
Other1	0.02	0.01	0.01	0.01	0.01	0.00	0.00	0.01	0.01	0.36	0.06	0.02	0.02	1.35	0.11	0.11	0.22	0.19	0.16	0.06	0.07	0.10	0.56	-5.33	1.46	0.17
Pioneer1	0.02	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.01	0.12	0.21	0.02	0.02	1.50	0.11	0.13	0.22	0.18	0.13	0.08	0.07	0.11	0.60	1.66	-5.57	0.17
Stine1	0.04	0.01	0.02	0.02	0.03	0.01	0.01	0.01	0.02	0.17	0.10	0.03	0.21	1.61	0.12	0.13	0.33	0.33	0.30	0.12	0.15	0.15	0.71	2.10	1.82	-8.26

substitute for AS1 is Asgrow0 (value of 2.50). The trait effect, however, dominates if we exclude the conventional Asgrow variety; the closest substitutes for AS1 are all non-Asgrow GT products. Finally, there is a strong assymetry in the cross-price elasticities between GT and conventional products of the same brand. For example, the cross-price elasticity of 0.07 from AS0 to Asgrow1 is a small fraction of the cross-price elasticity of 2.50 from AS1 to Asgrow0. This is simply due to the fact that GT products typically have much larger shares compared to the identically branded conventional versions (specifically, the denominator of equation (13) is much smaller for conventional products).

5.4 Reshuffling

Although our framework permits unobserved heterogeneity for all brands, we do restrict this heterogeneity to follow a normal distribution. Previous research has shown that, even having controlled for unobserved heterogeneity, it is still possible to incorrectly find positive evidence of structural state dependence if that unobserved heterogeneity is not sufficiently flexible enough. The term for this is spurious state dependence. To check whether unobserved heterogeneity has been captured in a sufficiently rich way, we conduct a reshuffling procedure along the lines of Dubé, Hitsch, & Rossi (2010) and Bronnenberg, Dubé, & Moorthy (2018). The basic idea of this procedure is to reshuflle the choice sequences in a random way and then re-estimate the full mixed logit model. This exploits the fact that structural state dependence should in principle only be identified by non-zero order features in the data, what some authors have referred to as "spells". If we have sufficiently controlled for unobserved heterogeneity, then the state dependent parameter should go to zero. On the other hand, if the estimate for γ remains large and positive, this may suggest that our original estimate is spurious and is likely due to an insufficiently rich accounting for unobserved heterogeneity.

For reshuffling, we first generate a new time variable whose values are drawn from a discrete uniform distribution with values ranging from 1996-2016. We then replace the original time variable with this new, randomly generated time variable. Consequently, we build a "reshuffled" purchase history for each farmer, which results in a new "randomized" state dependence term. In the process, three of the marketing variables are also reconstructed: ind_dealer, ind_rep, and ind_retailer. All other explanatory variables are held fixed throughout the reshuffle process. In other words, instead of randomly reshuffling the purchase sequence for each individual and then reconstructing new market-specific choice sets, we maintain the original choice set for each individiual and thereby *only* reshuffle the state dependence term. We do it this way for the following reasons. The main reason is that the industry has experienced signficant structural changes during the observed timespan—some brands have exited and entered the industry, prices have risen, and the GT trait has come to dominate. For these reasons, fully reconstructed choice-sets would have unreasonable properties. Consider the following example. Suppose an individual purchased the brand Channel with GT in 2015, and that upon reshuffling the new, randomly assigned year was 1996. If choice-sets were fully recontstructed, then Channel with GT will enter the choice set of all farmers in this local CRD in 1996 and its price will be an average price of all corresponding Channel products after reshuffle. This raises three problems: (i) Channel only enters the market after 2009; (ii) seed prices in the later periods are significantly higher than in early periods, even after deflation; and (iii) the size of any choice-set is subject to change in the reconstruction process.

The model results after reshuffling the data are presented in Table 9. We report results for two types of models: the basic conditional logit model, where unobserved heterogeneity is not captured by random coefficients, and the mixed logit model, which controls for unobserved heteroegeneity through the inclusion of random coefficients for the brand and trait dummies.

Overall, the coefficient for structural state dependence significantly decreases in both models, suggesting that state dependence is indeed a feature of soybean seed demand. In the conditional logit model, the coefficient decreases from 2.316 to 1.218, whereas in the mixed logit model, the coefficient decrease from 2.065 to 0.462. The smaller decrease in state dependence for the conditional logit model highlights the importance of controlling for unobserved heterogeneity.

Table 9. Model results after reshuffling the choice sequences

	Conditional logit	Mixed logit	
	Main	Main, coef	Main, SD
	b/se	b/se	
price	-0.181***	-0.264***	
	(0.014)	(0.024)	
state	1.218***	0.462***	0.339***
	(0.028)	(0.031)	(0.079)
GT 1999-2004	3.013***	4.696***	2.079***
	(0.210)	(0.331)	(0.111)
GT 2005-2010	4.437***	7.515***	2.687***
	(0.238)	(0.403)	(0.189)
GT 2011-2016	3.129***	6.833***	4.098***
	(0.184)	(0.393)	(0.318)
Asgrow	2.375***	4.008***	1.120***
	(0.413)	(0.683)	(0.050)
Beck	3.822***	4.380***	2.751**
	(0.469)	(1.166)	(1.040)
Channel	2.702***	1.575*	3.142***
	(0.408)	(0.745)	(0.201)
Croplan	3.056***	4.262***	-1.252***
	(0.421)	(0.678)	(0.132)
DeKalb	0.133	1.295*	0.634***
	(0.383)	(0.623)	(0.168)
Golden	0.087	0.653	1.362***
	(0.465)	(0.701)	(0.104)
Growmark	2.045***	3.160***	1.581***
	(0.432)	(0.665)	(0.269)
Kruger	1.703***	1.683*	2.171***
	(0.458)	(0.788)	(0.181)
Mycogen	2.011***	2.778***	1.782***
	(0.398)	(0.715)	(0.171)
NK	2.544***	4.039***	1.149***
	(0.381)	(0.625)	(0.107)
Other	2.569***	4.006***	1.159***
	(0.363)	(0.601)	(0.089)
Pioneer	2.597***	4.274***	0.891***
	(0.379)	(0.637)	(0.128)
Stine	2.119***	3.484***	0.985***
	(0.430)	(0.728)	(0.288)
control	0.113***	0.169***	0.170***
	(0.013)	(0.025)	(0.007)
LL	-141776.678	-117337.689	
N	1057637	1057637	

Note: standard errors are clustered at CRD level.
*p<0.05, **p<0.01, ***p<0.001. The sign of the estimated standard deviations is irrelevant: consider them as positive.

While the state dependence coefficient does decrease signficantly in both cases, particularly in the full model, it still remains positive with a coefficient of 0.462. This may suggest that the assumption of normally distirbuted random coefficients is not rich enough and therefore a portion of our original state dependence estimate is spurious. We note, however, that an alternative possiblity is that this is the result of a higher-order Markov chain (more distant purchases may still have some impact on farmers' seed choice), or even the result of our large sample size. In any case, our reshuffling procedure suggests that at least 80% of what we captured in the mixed logit model in Table 6 is the result of genuine structural state dependence. To the extent that there is some bias in the structural state dependence coefficient, it is small (about $1.6/unit).

6. Counterfactual analysis

Over the last three decades, the seed markets have been characterized by significant technological innovation (in the form of GE traits) and changing market structure (e.g., higher concentration). In this section, we assess some potential implications of state dependence, especially as it pertains to the introduction and diffusion of the GT technology.

In section 6.1, we explore whether state dependence affected the diffusion of the GT technology. In particular, we simulate GT adoption rates in two scenarios: one with state dependence and the other without state dependence. In section 6.2, we examine how state dependence and the GT technology contributed to changing brand shares in the industry. To do so, we simulate brand market shares in four settings: (i) an environment with both the GT trait and state dependence; (ii) an industry without the GT trait but still with state dependence; (iii) an industry with the GT trait but without state dependence; and (iv) an industry with neither GT nor state dependence. We are particularly interested in whether state dependence conferred an "early mover" advantage for brands that were the first to embed and offer the GT trait in their seed varieties. This is similar to the question of whether there is an advantage to being the first brand to locate in a particular geographic region (Bronnenberg, Dhar, and Dubé 2009; Bronnenberg, Dubé, and Gentzkow 2012), however, our exploration assesses whether state dependence confers an advantage to those brands first to locate in a new dimension of the product space.

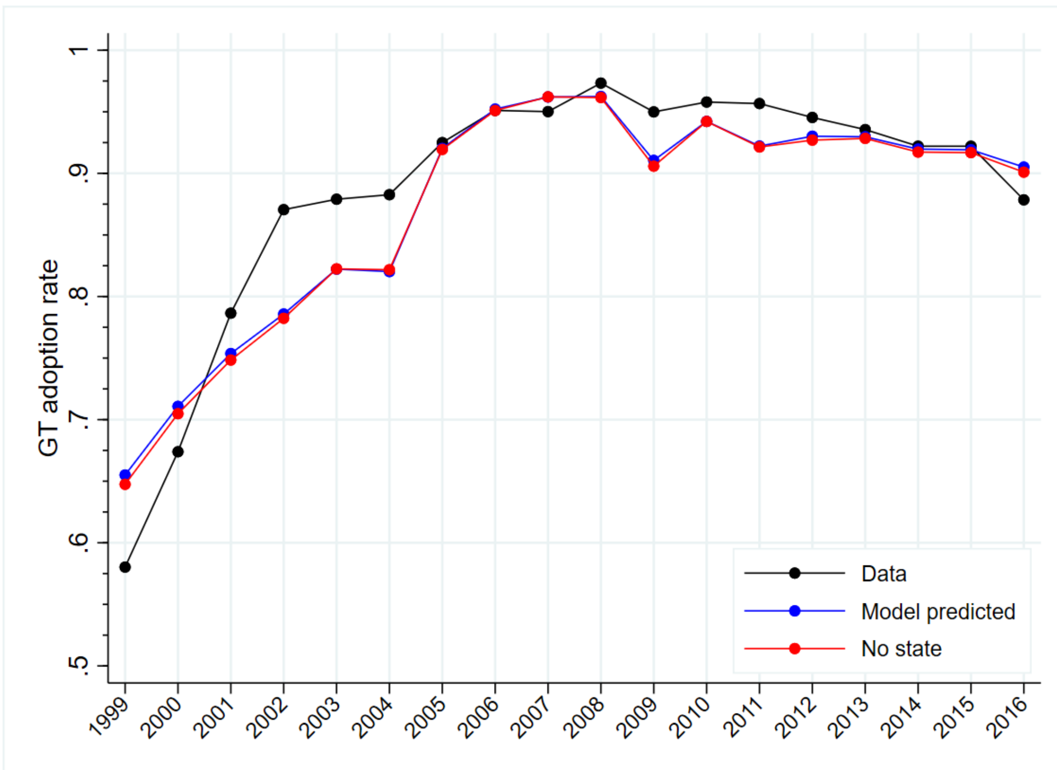
6.1 GT adoption rate

To assess the importance of structural state dependence for the diffusion of a new major technology, in this case GT adoption, we calculate predicted GT adoption rates under two different scenarios: in the first scenario (or baseline scenario), we simply generate predicted GT adoption rates using the estimated coefficients from the full mixed logit model. In the counterfactual scenario, we compute predicted GT adoption rates under the assumption that there is no structural state dependence; i.e., when $\gamma^h = 0$. For comparison, we also present the observed adoption rate ("original data"). As was done for the computation of elasticities, we first compute predicted individual demands and then aggregate those demands to national level using the observed amount of purchased units as weights. The GT adoption rate is calculated by summing over the predicted shares of products that contain GT. However, unlike in section 5.3, where we used the *observed* purchase history for the state dependence variable, here we reconstruct the state dependence variable using the predicted purchase probabilities from the most recent period. See the Appendix for further discussion and details on how the predictions were generated.

Figure 3 presents GT adoption rates for each case; black represents the observed rate, blue for the baseline model prediction, and red for the counterfactual scenario without state dependence. The results show that our model does fairly good job of predicting the GT adoption rate (the black line and the blue line). However, the model is least accurate in the first period (1999-2004). This likely has to do with how we have specified the GT-period interaction variables. In the first period (1999-2004), the coefficient on the GT-period interaction variable is fixed. Because actual adoption rates changed significantly during this time, there is over-estimation in the first couple years (1999-2000) and an under-estimation in the last few years (2002-2004).

By comparing the scenario without state dependence to the scenario with state dependence, we see that state dependence has little effect on the diffusion of GT varieties. There is a very small positive effect in the first few years (the red line is below the blue line before 2002), but the effect is negligible in later years. In some ways, this isn't terribly surprising. In order for state dependence to impact the GT adoption rate it would need to be the case that by removing state dependence there is a systemic shift in preferences towards either varieties with GT (or without). In our case, this should happen through differentiated GT products of different brands (some brands have GT products, some do not). However, because of aggressive licensing of GT traits to other seed suppliers by Monsanto, most brands in the industry added GT to their varieties very quickly.

Figure 3. GT adoption rates over time



6.2 Brand market shares

Using the full model estimates, we simulate brand market shares in four settings: (i) an environment with both the GT trait and state dependence; (ii) an industry without the GT trait but still with state dependence; (iii) an industry with the GT trait but without state dependence; and (iv) an industry with neither GT nor state dependence. As done previously, we first compute predicted individual demands in each scenario and then aggregate to the national level for each brand using projected purchased units as weights.

In simulating scenarios without the GT trait, there is an important issue that needs to be resolved. One way to simulate the non-GT counterfactual would be to remove GT products from farmers' choice sets. However, because GT products had almost completely crowded out non-GT products by 2007, this would artificially reduce farmers' options. Therefore, we construct the non-GT counterfactual in a more cautious way—we remove the GT attribute from a product (rather than remove the product) and maintain the newly created "synthetic" conventional product if there was no pre-existing conventional version. For example, consider the brand Channel, which was only sold with the GT trait. Under this counterfactual, the product Channel1 becomes Channel0.

In conducting this exercise an additional issue that needs to be addressed is what the prices for these new alternatives would have been in the counterfactual. As noted in CIliberto, Moschini, & Perry (2019), a structural equilibrium model may be ideal but difficult to execute in the seed market because of cross-licensing for GT traits. Thus, we follow the approach of Hausman & Leonard (2002) by using a hedonic price regression to construct the counterfactual prices.

6.2.1 Hedonic price regression

As discussed previously, we construct a contingent price system to account for seasonal discounts in seed prices. Consequently, we run two hedonic regressions, one for market retail prices, and one for market net prices, and then assign the corresponding prices to each choice set, contingent on whether a discount was received. The hedonic regression is written as follows:

$$p_{jm} = \phi' \tilde{x}_{jm} + \theta D_m + \xi_t + \xi_l + e_{jm} \tag{13}$$

We estimate this regression at the market level—there are 21,008 observations across 2,609 markets, with roughly 8 alternatives per market. The vector $\tilde{x}_{jm}$ contains all variables in x_{jm} except the initial conditions and the interactions of marketing variables with individual purchasing experiences. To capture any potential competitive price effects from the introduction of GT products, we add D_m as in CIliberto, Moschini, and Perry (2019), which is an indicator variable of whether GT product is purchased in the market. The last two variables ξ_t and ξ_l are time fixed effects and CRD fixed effects, respectively, where t and l are both identified by market m. Finally, e_{jm} is an idiosyncratic shock, which is assumed to be exogenous with all explanatory variables.

The hedonic regression results are presented in Table 10, which contains estimated coefficients for the GT-period interaction variables, the brand intercepts (in the last period), and the indicator variable for the presence of GT products. These results convey the average price premiums that seed companies charge for GT varieties. Generally, GT seeds are associated with higher prices, with the premium being the smallest in the final period. Our estimates are very close with the results from a more comprehensive hedonic regression in CIliberto, Moschini, & Perry (2019).[28] We see a drop in the price premium for GT trait in the last period, which is in line with the WTP estimates in Table 7. Comparing the regression of retail prices and that of net prices, it shows that the price premium is lower for GT trait in the last two periods, which may imply that GT products are generally associated with more discounts in those periods. Furthermore, note that the estimates here are smaller than the WTP estimates in Table 7, suggesting the potential welfare gains of farmers during the introduction of this new technology (Ciliberto, Moschini, and Perry, 2019)

Finally, the coefficient of the GT indicator variable is insignificant, which may be ascribed to a lack of observations. As our regression start price 1999, when the GT adoption rate is over 50%, only 2 markets with 5 observations are absent of GT products: the variable D_m can hardly be identified in our case.

[28] In CIliberto, Moschini, & Perry (2019), they estimate hedonic regressions for both soybean seeds and corn seeds. Adjusted by unit/acre ratio, they find that the price premium is 12.993 $/unit for soybean GT trait in 1996-2000, 11.660 $/unit for soybean GT trait in 2001-2006, and 9.627 $/unit for soybean GT trait in 2007-2011.

Table 10. Hedonic prices

	Retail b/se	Net b/se
GT 1999-2004	10.906***	10.717***
	(0.290)	(0.261)
GT 2005-2010	11.731***	10.643***
	(0.313)	(0.281)
GT 2011-2016	7.391***	5.987***
	(0.300)	(0.269)
D	3.542	3.662
	(2.275)	(2.042)
R2	0.822	0.807
N	21008	21008

Note: *p<0.05, **p<0.01, ***p<0.001

To construct the counterfactual prices, we set the indicators of trait to zero, which consist of GT trait variables in all three period as well as the interactions of GT with each brand. We then use the model estimates to predict counterfactual retail prices and net price, and further assign corresponding predicted prices to counterfactual choice sets contingent on whether the original purchase is offered a discount.

6.2.2 Brand market shares in four scenarios

The following Figure 4 depicts market shares of the two largest brands, Pioneer and Asgrow, in three of the four scenarios—with both state dependence and GT (blue), without state dependence but with GT (red), without GT but with state dependence (orange). To check our model fit, as in section 6.1, we add a black line to represent the observed market shares from the original data.

Comparing the blue line and the black line, we see that our model can capture the general trend and the magnitude of the true market shares but the model is not good at matching the year-by-year variance, which may be related to our classification of the three periods.

Figure 4. Market shares of Asgrow and Pioneer in three of the four scenarios

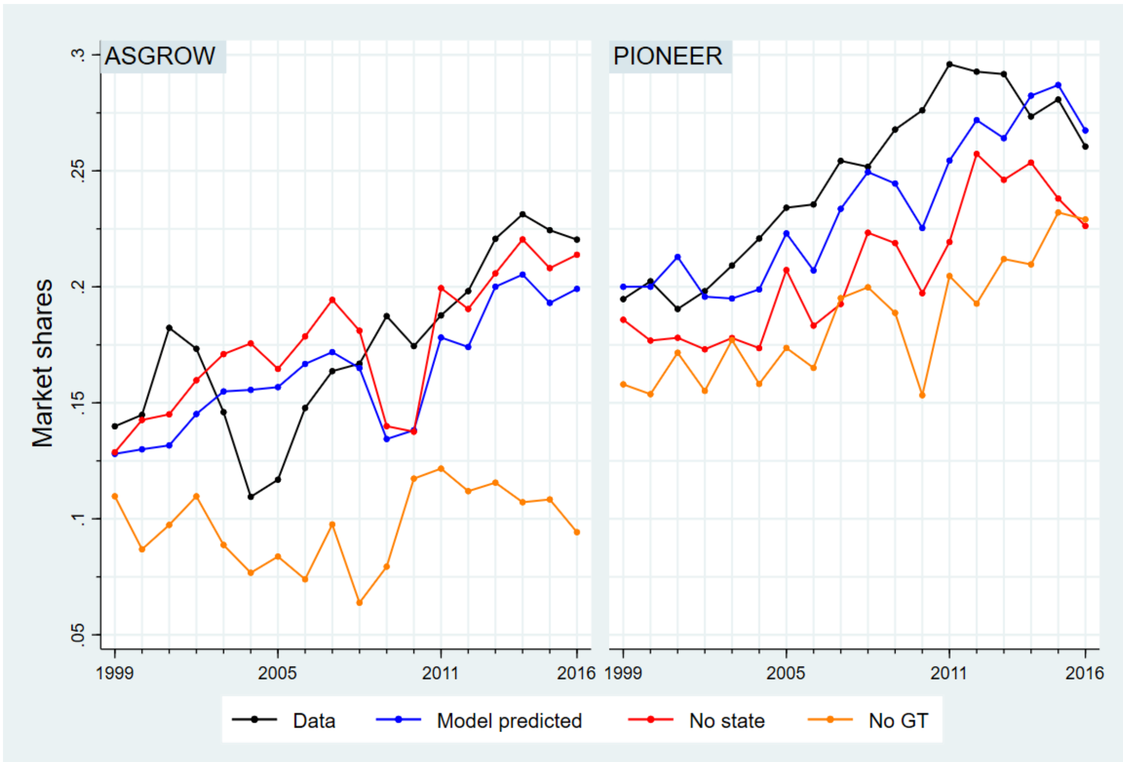
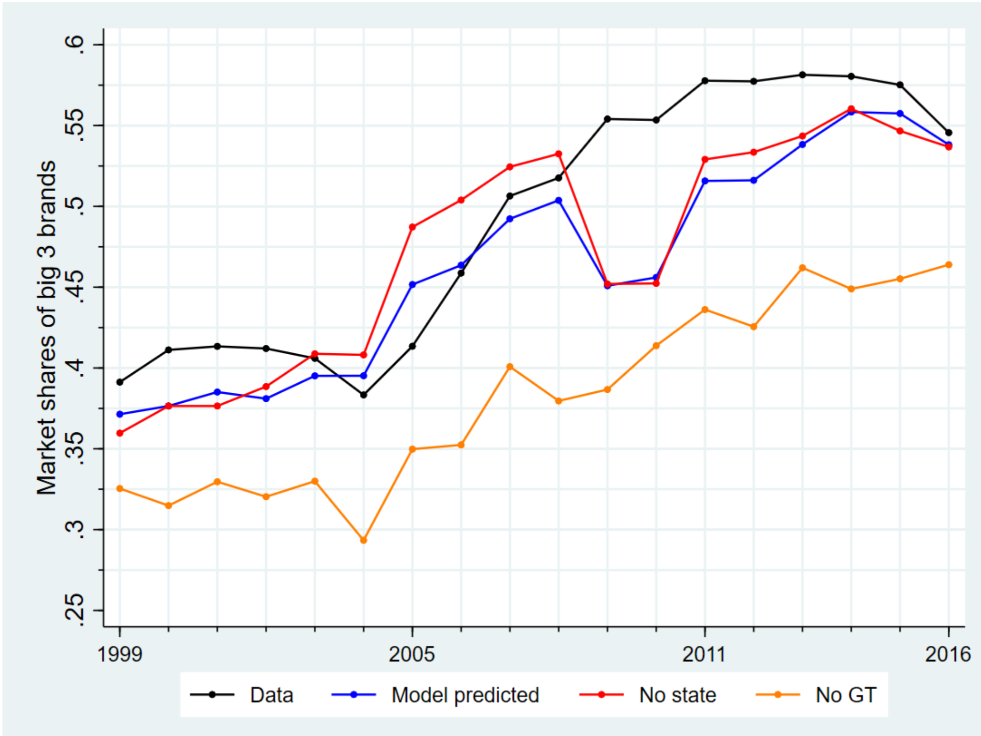


Figure 4 also shows that state dependence and the GT trait have different implications for Asgrow and Pioneer. The GT trait benefits both brands, but to different degrees: GT confers a larger increase in shares for Asgrow than for Pioneer, especially in the final period. Each brand is also affected by the presence of state dependence, but here the effects go in the opposite direction. Whereas Pioneer is benefitted by state dependence, Asgrow is slightly hurt by state dependence. We speculate that these differences are the result of differences in market shares: intuitively, larger brands benefit more from state dependence, as more farmers are willing to pay a premium to purchase them. Indeed, for the remaining brands, all smaller than Asgrow and Pioneer, state dependence has a detrimental impact on market share.

Figure 5. Simulated market shares for the big 3 brands (Asgrow, NK, Pioneer)

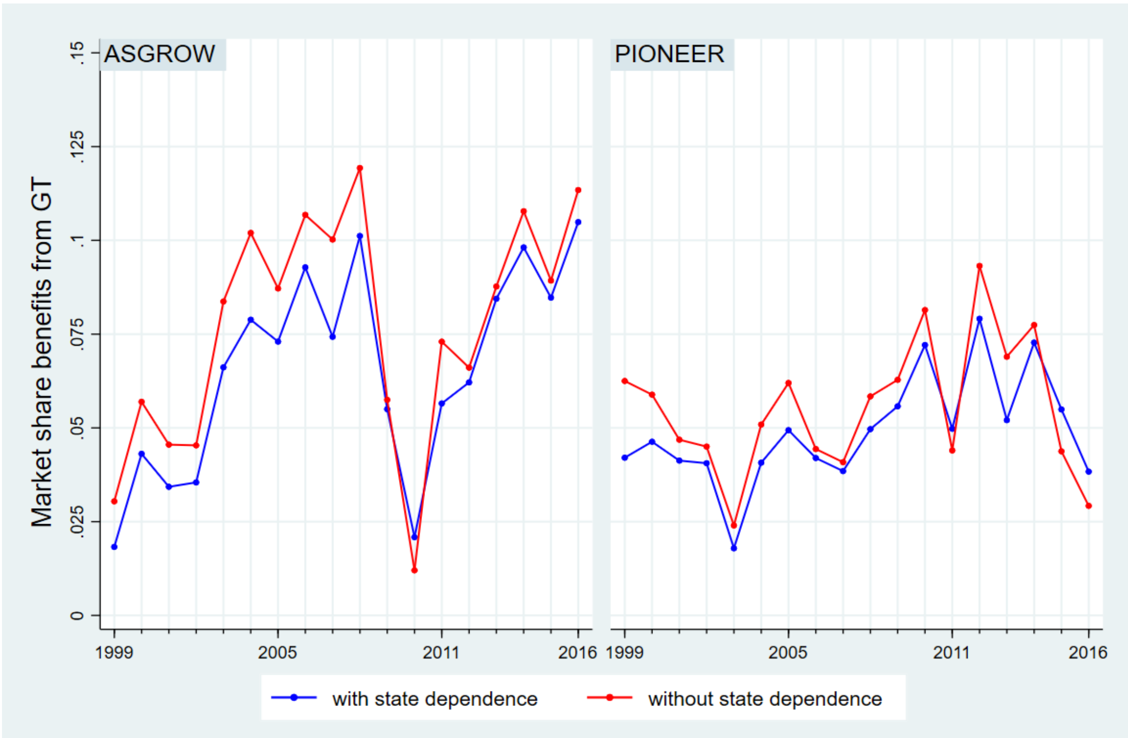


We next examine the implications of the GT trait and state dependence for market concentration. Figure 5 contains total aggregate market share for the big 3 brands as of the final period: Pioneer, Asgrow, and NK. We first note that the market is increasingly concentrated over time, with the three big three obtaining more than 50% of the market by the final period. Second, the GT trait was also a significant contributor to concentration, raising the big three share by about 10%. Third, state dependence does not have a significant effect on market concentration. This appears to be the result of the benefits for Pioneer being cancelled out by the losses for Asgrow and NK.

6.2.3 Structural benefits from GT

Although Figure 4 clearly depicts the impacts of removing the GT trait for Asgrow and Pioneer, it does not provide clarity on the question of whether structural state dependence affects the market share impacts of the GT trait. To explore this question, Figure 6 contains market share impacts of the GT trait with and without state dependence for Asgrow and Pioneer. Specifically, the blue line represents the gain in market share from the GT trait when state dependence is present, which we denote by GT_state_{bt}; and the red line represents the change in market share from the GT trait when state dependence is absent, denoted by $GT\_nostate_{bt}$.

Figure 6. Market share benefits from GT



By comparing the blue line and the red one, we see that state dependence has actually reduced the benefits of the GT trait for Asgrow and Pioneer. This may have something to do with changes in shares over time; generally, when share increase, state dependence seems to have a negative effect.

Figure 6 demonstrates the impact of the GT trait and state dependence for the entire sample period. However, state dependence may exert the largest influence in the early stages of adoption through an "early mover" effect. Conceptually, if a brand is one of the first to offer a new technology, in this case the GT trait, it can attract new customers. Throught state dependence, it can obtain an extra advantage relative to brand that add the GT trait later on. In in the industry, Asgrow, NK, and Pioneer were all early adopters and also the three largest brand at the end of our timespan. To investigate the issue of whether there was an early mover advantage, we regress GT_state_{bt} on early adoption rates, denoted by r_b, for each brand. The early adoption rate is defined as the product-specific average market share of the GT trait in the first two years, 1996 and 1997.[29] In the regression, we discard observations for "PUBLIC" and "CHANNEL", so there are 216=18 (number of years) *12 (=14-2, number of considered brands) total observations. The regression we estimate is written as

$$GT\_state_{bt} = ar_b + u_{bt}, \tag{14}$$

where u_{bt} is a brand-time-specific error term. The results are shown in the first column of Table 10.

The coefficient of r_b is both positive and significant, along with a high R^2, suggesting that a brand's structural benefits from the GT technology can largely be explained by the degree to which it was an early seller of the technology. However, this result does not convey the impact of state dependence this benefit. Therefore, we estimate a second specification that conditions on the GT impact that is not due to state dependence (denoted by $GT\_nostate_{bt}$). This variable controls for other brand-specific factors related to early adoption, such as more varieties with the GT trait or a better distribution channel. This regression is written as:

$$GT\_state_{bt} = ar_b + bGT\_Nostate_{bt} + \tilde{u}_{bt} \tag{15}$$

Table 10. How brands benefit from early adoption of GT seeds

	reg1 b/se	reg2 b/se
r_b	2.867***	0.357***
	(0.126)	(0.081)
GT_state_{bt}		0.773***
		(0.020)
R2	0.708	0.962
N	216	216

Note: *p<0.05, **p<0.01, ***p<0.001

[29] The regression is estimated on the sample period 1999-2016. This is done to create a gap between the time when the early adoption shares are computed (1996-1997) and when the GT share impacts are regressed.

The results are reported in the second column "reg2" of Table 10. After controlling for benefits of GT technology without state dependence, we see that the coefficient of r_b has a sharp drop, suggesting that the early adoption of the GT technology benefits the brand mainly through the brand-specific factors. However, the coefficient of r_b is still positive and significant, suggesting the brand also benefits from early adoption through structural state dependence.

7. Conclusion

In this paper, we develop and estimate a micro-level structural model of U.S. soybean seed demand to study a recurring theme in economics and marketing, brand inertia. To disentangle unobserved heterogeneity from state dependence, we adopt the random coefficients logit model for demand estimation. To deal with the initial conditions problem, we use the initial brand choices as an extra explanatory variables. To deal with price endogeneity, we apply the control function approach, using the previous year's futures price as a cost instrument and interacting them with brand and GT trait dummies.

Our results show that structural state dependence generally exists for all farmers, with an average WTP of \$6.77/unit, which is about 15% of the average soybean retail price of \$45/unit. We also find that state dependence is quite heterogeneous—state dependence is valued at more than \$10/unit for 15% of farmers, whereas another 15% values it at less than \$4/unit. Along with demand estimation, we show that farmers' WTPs for the GT trait vary over time and over individuals. On average, the WTP for GT is \$17.09/unit during 1999-2004, goes up to \$27.71/unit during 2005-2010, and then declines to \$22.89/unit during 2011-2016. Adjusted by unit/acre ratio, our results are consistent with the WTP estimates reported by Ciliberto, Moschini, and Perry (2019), who employ a different modeling approach. Our finding of declining WTP for the GT trait in the 2011-2016 period is also consistent with the emergence of glyphosate resistant weed (Perry, Ciliberto, Hennessy, and Moschini, 2016). We further show that farmers are quite heterogeneous towards GT traits, with much higher variance than the state dependence coefficient.

Using the estimated model, we generate own and cross-price demand elasticities for each product. We find that the own-price elasticities are, on average, -5.4 for conventional products and -7 for GT products. Generally, farmers are more likely to substitute among products of the same brand and same trait, however, the strength of these effects differ depending on the type product. In particular, a farmer is more likely to switch to another conventional product for a different brand if she chooses a conventional product, whereas she is more likely to switch to the conventional product of the same brand if she chooses a GT product.

Finally, we assess some potential implications of state dependence during the introduction and diffusion of the GT technology. We obtain four main findings. First, we find that state dependence has almost no effect on the diffusion of GT varieties, a reflection of the fact that most brands in the industry added the GT trait to their varieties very quickly. Second, we show that brands benefit differently from state dependence. The largest brand, Pioneer, benefits from state dependence whereas the other, smaller brands are negatively affected. Next, concerning the impact of GT on brand market shares, we show that state dependence acts as a cushion—it pulls the effect of GT trait diffusion, whether positive or negative, back to zero. Finally, we examine whether state dependence confers an advantage to early providers of the new technology. The result is small but significantly positive: brands benefit from early adoption, even after controlling for the benefits of GT when state dependence is absent.

Reference

Akay, A. (2012). Finite-sample comparison of alternative methods for estimating dynamic panel data models. Journal of Applied Econometrics, 27(7), 1189-1204.

Arulampalam, W., & Stewart, M. B. (2009). Simplified implementation of the Heckman estimator of the dynamic probit model and a comparison with alternative estimators. Oxford bulletin of economics and statistics, 71(5), 659-681.

Barrows, G., Sexton, S. and Zilberman, D. (2014) Agricultural Biotechnology: The Promise and Prospects of Genetically Modified Crops. Journal of Economic Perspectives, Vol. 28, pp. 99-119.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. The RAND Journal of Economics, 242-262.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. Econometrica: Journal of the Econometric Society, 841-890.

Berto Villas-Boas, S. (2007). Vertical relationships between manufacturers and retailers: Inference with limited data. The Review of Economic Studies, 74(2), 625-652.

Bronnenberg, B.J., Dhar, S.K. and Dubé, J.P.H. (2009). Brand history, geography, and the persistence of brand shares. Journal of political Economy, 117(1), pp.87-115.

Bronnenberg, B. J., Dubé, J. P. H., & Gentzkow, M. (2012). The evolution of brand preferences: Evidence from consumer migration. American Economic Review, 102(6), 2472-2508.

Bronnenberg, B. J., Dubé, J.-P., & Moorthy, S. (2019). The Economics of Brands and Branding. In: Dubé, J. P. & Rossi, P. E., eds, Handbook of the Economics of Marketing, Elsevier B.V., chapter 6, pp. 291-358.

Cameron, A.C. and Trivedi, P.K. (2005). Microeconometrics: methods and applications. Cambridge university press.

CIliberto, F., Moschini, G., & Perry, E. (2019). Valuing Product Innovation: Genetically Engineered Varieties in US Corn and Soybeans. The RAND Journal of Economics, 50(3), 615-644.

Clancy, M. S., & Moschini, G. (2017). Intellectual property rights and the ascent of proprietary innovation in agriculture. Annual Review of Resource Economics, 9, 53-74.

Dubé, J. P., Hitsch, G. J., & Rossi, P. E. (2010). State dependence and alternative explanations for consumer inertia. The RAND Journal of Economics, 41(3), 417-445.

Dubé, J.-P., Hitsch, G. J., & Rossi, P. E. (2009). Do switching costs make markets less competitive? Journal of Marketing research, 46(4), 435-445.

Fernandez-Cornejo, J., & Spielman, D. J. (2002). Concentration, market power, and cost efficiency in the corn seed industry, working paper.

Fernandez-Cornejo, J. (2004). The Seed Industry In U.S. Agriculture: An Exploration Of Data And Information On Crop Seed Markets, Regulation, Industry Structure, And Research And Development. United States Department of Agriculture, Economic Research Service.

Fernandez-Cornejo, J., Hendricks, C., & Mishra, A. (2005). Technology Adoption and Off-Farm Household Income: The Case of Herbicide-Tolerant Soybeans. Journal of Agricultural and Applied Economics, 37(3), 549-563.

Funk, T. F., & Vincent, A. T. (1978). The farmer decision process in purchasing corn herbicides. No. 1620-2016-134790.

Goldberg, P. (1995). Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry. Econometrica, 63(4), 891-951.

Graff, G. D., Rausser, G. C., & Small, A. A. (2003). Agricultural biotechnology's complementary intellectual assets. Review of economics and statistics, 85(2), 349-363.

Handel, B. R. (2013). Adverse selection and inertia in health insurance markets: When nudging hurts. American Economic Review, 103(7), 2643-82.

Harbor, A. L., Martin, M. A., & Akridge, J. T. (2008). Assessing input brand loyalty among US agricultural producers. International Food and Agribusiness Management Review, 11(1030-2016-82702), 17-34.

Hausman, J. A., & Leonard, G. K. (2002). The competitive effects of a new product introduction: A case study. The Journal of Industrial Economics, 50(3), 237-263.

Heckman, J.J. (1981). Heterogeneity and state dependence. In Studies in labor markets (pp. 91-140). University of Chicago Press.

Heckman, J. J. (1987). The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process and some Monte Carlo evidence (pp. pp-114). University of Chicago Center for Mathematical studies in Business and Economics.

Hole, A. R. (2007). Fitting mixed logit models by using maximum simulated likelihood. The Stata Journal, 7(3), 388-401.

Horsky, D., Misra, S., & Nelson, P. (2006). Observed and unobserved preference heterogeneity in brand-choice models. Marketing Science, 25(4), 322-335.

ISAAA. (2016). Global Status of Commercialized Biotech/GM Crops: 2016.

Keane, M. P. (1997). Modeling heterogeneity and state dependence in consumer choice behavior. Journal of Business & Economic Statistics, 15(3), 310-327.

Kim, H., & Moschini, G. (2018). The dynamics of supply: US corn and soybeans in the biofuel era. Land Economics, 94(4), 593-613.

Kohls, R. L., Stucky, R. L., & Gifford, J. I. (1957). Farmers' Selection of Farm Machinery Dealers. Journal of Marketing, 21(4), 446-450.

Kool, M. (1994). Vendor loyalty of farmers: Characterisation, description and analysis. European Review of Agricultural Economics, 21(2), 287-307.

Lamkey, K. (2004). Seed production in corn and soybean. Agronomy Reports. 4. https://lib.dr.iastate.edu/agron_reports/4/

Louviere, J., Train, K., Ben-Akiva, M., Bhat, C., Cameron, T. A., Carson, R. T., . . . Waldman, D. (2005). Recent progress on endogeneity in choice modeling. Marketing Letters, 16(3-4), 255-265.

MacKay, A. and Remer, M., 2019. Consumer Inertia and Market Power. Available at SSRN 3380390.

Moschini, G. (2008). Biotechnology and the development of food markets: retrospect and prospects. European Review of Agricultural Economics, 35(3), 331-355.

Moschini, G. (2010). Competition issues in the seed industry and the role of intellectual property. Choices, 25(2), 1-12.

Nevo, A. (2000). A practitioner's guide to estimation of random-coefficients logit models of demand. Journal of economics & management strategy, 9(4), 513-548.

OECD. (2018). Concentration in Seed markets: Potential Effects and Policy Responses. Paris: OECD.

Perry, E. D., Ciliberto, F., Hennessy, D. A., & Moschini, G. (2016). Genetically engineered crops and pesticide use in US maize and soybeans. Science advances, 2(8), e1600850.

Perry, E.D., Moschini, G. and Hennessy, D.A., 2016. Testing for complementarity: Glyphosate tolerant soybeans and conservation tillage. American Journal of Agricultural Economics, 98(3), pp.765-784.

Perry, E.D., Hennessy, D.A. and Moschini, G., 2019. Product concentration and usage: Behavioral effects in the glyphosate market. Journal of Economic Behavior & Organization, 158, pp.543-559.

Petrin, A., & Train, K. (2010). A control function approach to endogeneity in consumer choice models. Journal of marketing research, 47(1), 3-13.

Schenkelaars, P., de Vriend, H., Kalaitzandonakes, N., Magnier, A., & Miller, D. (2011). Drivers of Consolidation in the Seed Industry and its Consequences for Innovation. Report commissioned by COGEM.

Seetharaman, P. B. (2004). Modeling multiple sources of state dependence in random utility models: A distributed lag approach. Marketing Science, 23(2), 263-271.

Seetharaman, P. B., & Chintagunta, P. (1998). A model of inertia and variety-seeking with marketing variables. International Journal of Research in Marketing, 15(1), 1-17.

Sellars, S. C., & Gunderson, M. A. (2018). Brand and Dealer/Retailer Loyalty among Large U.S. Farmers. Selected presentation in 2018 AAEA.

Shi, G., Chavas, J.-P., & Stiegert, K. (2010). An analysis of the pricing of traits in the US corn seed market. American Journal of Agricultural Economics, 92(5), 1324-1338.

Simonov, A., Dubé, J.P., Hitsch, G.J., and Ross, P. (2019) State-dependent demand estimation with initial conditions correction. NBER Working Paper No. 26217, September.

Sudhir, K., & Yang, N. (2014). Exploiting the Choice-Consumption Mismatch: A New Approach to Disentangle State Dependence and Heterogeneity. Working paper.

Syngenta. (2016). Our industry 2016. Retrieved from: https://www.syngenta.com/~/media/Files/S/Syngenta/documents/our-industry-syngenta.pdf

Train, K. E. (2009). Discrete choice methods with simulation. Cambridge university press.

Train, K. E., & Winston, C. (2007). Vehicle choice behavior and the declining market share of US automakers. International economic review, 48(4), 1469-1496.

Train, K., & Weeks, M. (2005). Discrete choice models in preference space and willingness-to-pay space. Applications of simulation methods in environmental and resource economics, 1-16.

Wechsler, S. J., McFadden, J. R., & Smith, D. J. (2018). What do farmers' weed control decisions imply about glyphosate resistance? Evidence from surveys of US corn fields. Pest management science, 74(5), 1143-1154.

Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. Journal of Applied Econometrics, 20(1), 39-54.

Wooldridge, J. M. (2015). Control function methods in applied econometrics. Journal of Human Resources, 50(2), 420-445.