GM vs. Non-GM: A Survival Analysis of U.S. Hybrid Seed Corn

Xingliang Ma and Guanming Shi

We analyze the rate of survival of hybrid seed corn in the U.S. market between 2000 and 2007 and find that product characteristics and market structures have significant impacts. The results suggest that survival rates depend on spillover effects of learning regarding genetically modified seeds and close substitute seeds and that hybrids planted in the Corn Belt survive longer in the market than hybrids planted in other parts of the country. Given increasing concentration in seed and agricultural biotechnology sectors, our results shed light on factors that affect the survival of hybrid corn seeds under various market conditions.

Key Words: GM technology, market structure, survival analysis, United States hybrid seed corn

The development of agricultural seeds through breeding programs represents a critical step in producing crops in virtually every region of the world. Losing a comparative advantage in producing specific crops usually means that farmers need to exit that market or obtain government subsidies. A common phrase in United States policy, “corn is king,”1 refers to the dominant role corn plays in agriculture in the United States as a feed grain and the significant government support the industry receives through direct subsidies, ethanol subsidies, and sugar quotas. While the U.S. government has clearly incentivized corn production, development of corn hybrids has largely been a private-sector enterprise. Before the 1990s, there was not much in the way of public policy concerning the strategic behavior of firms that develop and sell corn hybrid seeds; the industry was not overly concentrated and yields of corn rose steadily in the United States.

The emergence in the 1990s of patented genetic modifications to seeds (together with changes to the mechanisms that protect intellectual property rights associated with plants) has been a catalyst for structural changes in agricultural seed sectors, including corn, soybeans, and cotton (Fernandez-Cornejo 2004). These changes included acquisitions of small firms that held patents on promising traits, vertical mergers between large biotechnology firms and local or regional seed companies, and the exit or horizontal merging of seed

1 The documentary “King Corn” helped to spread the popular concept of corn’s dominance as a row crop in the United States.
companies (e.g., Fernandez-Cornejo 2004, Shi 2009, Shi, Chavas, and Stiegert 2010). Such rapid restructuring of the U.S. seed sector has drawn the attention of academic researchers, policymakers in Washington D.C., and various farm groups. At issue is not only the critical concern about seed pricing but also the availability of seeds as the market becomes more concentrated. Is restructuring of the seed industry changing the basic nature of seed survival rates? Will declining competition in the seed market lead to undesirable outcomes for corn farmers who are looking for the best seeds for their specific operations?

Given rapidly changing technology and many ongoing mergers and acquisitions in the industry, it can be difficult to find conclusive evidence of the effects of these changes in market structure. We aim to find some middle ground where, in spite of the empirical difficulties, an analysis of the survival of hybrid seeds in the corn seed market can generate insights that help answer such research and policy questions. Survival analyses are commonly referred to as “product life cycle” analyses and provide useful information for firms trying to balance research and development (R&D) expenditures with marketing and product portfolio investments. A firm naturally prefers to remain competitive in a market while not spending much on developing new products. When an industry is highly concentrated, the need to emphasize new product development may not be urgent. However, failing to innovate while other firms do so can have disastrous consequences; product survival can have an impact on survival of the firm (e.g., Klepper 1996, Agarwal and Gort 2002).

The empirical literature on product survival is somewhat limited. Bayus (1998) investigated whether a product life cycle shrinks as technology progresses using data on mainframe computers that were introduced in the United States between 1974 and 1992. He found no evidence of reductions in product model life. Both Greenstein and Wade (1998) and Asplund and Sandin (1999) examined the impact of imperfect market competition on the life of a product. We might expect that a firm that faces limited competition would enjoy longer life for its products since there are relatively few alternatives available to consumers. Using data for 1968 through 1982, Greenstein and Wade (1998) found that the survival rate of commercial mainframe computers tended to decrease when the number of competing firms increased. De Figueiredo and Kyle (2001, 2006) examined the market for laser printers and also found that competition had a large impact, accelerating product exits and delaying product entries. They found that firms with strong brands and high market shares could keep their products in the market longer and introduced few products. However, Asplund and Sandin (1999) examined the Swedish market for beer from 1989 to 1995 and found that products introduced by firms that held a large share of the market did not survive as long as brands produced by other firms, potentially due to the effects of product cannibalization within the large firms. Indeed, Pinkse and Slade (2004), in an examination of the market

---

2 In May 2009, Assistant Attorney General Christine A. Varney (2009) signaled a renewed interest at the U.S. Department of Justice (DOJ) in vigorous anti-trust enforcement. In December 2009, the U.S. Department of Agriculture (USDA) and the Farm Foundation hosted a conference on competition issues in agriculture that included presentations on the seed industry. Throughout 2010, USDA and DOJ held joint regional meetings at various locations around the United States and one in Washington D.C. to discuss competition issues that included conditions in the seed industry.

3 Most of the existing literature on life cycle analysis has focused on survival at the firm level, such as a firm’s entry and/or exit decisions (e.g., Klepper and Simons 2000, Fontana and Nesta 2009, Nikolaeva et al. 2009).

4 Multi-product firms may encounter internal competition between their various products,
for beer in the United Kingdom, found that changing a market’s structure (via mergers) could affect brand disappearance and introduction (the product space).

A product’s survival in a market may depend on its ability to compete against substitute products both horizontally (e.g., in terms of differentiated tastes) and vertically (e.g., in terms of quality). Therefore, characteristics or attributes of a product play a critical role in determining how long the product can remain in the market. Requena-Silvente and Walker (2009) investigated the duration of differentiated products in the automobile market in the United Kingdom between 1971 and 2002 and focused on the quality and characteristics of the products with controls for market structure. They found that the initial location in the product-characteristic space and repositioning of that location through introduction of new variants, together with quality upgrades, significantly extended the life cycle of a car model.

Our study also focuses on the survival of a differentiated product, genetically modified (GM) and conventional corn seeds. Magnier, Kalaitzandonakes, and Miller (MKM) (2010) studied the U.S. hybrid seed corn market, focusing on product life cycles of GM and conventional seeds and the role of innovation (i.e., biotechnology). They demonstrated that changes in biotechnology are important factors that influence seed survival. Our analysis diverges from MKM’s study by jointly estimating impacts on the survival rates of corn seeds from (i) various GM characteristics embedded in seeds and (ii) the rapidly changing horizontal and vertical market conditions under which the seeds are sold. Not only have GM seeds taken a dominant market position in corn production, but the corn hybrid seed market has become more concentrated and vertically integrated with the upstream biotechnology sector. More than 200 corn seed companies operated in various local markets in the United States in 2007. Only four of those firms were vertically integrated with the upstream biotechnology sector, but the seeds those companies sold have generated around 70 percent of the nation’s total corn acreage since 2005 (Stiegert, Shi, and Chavas 2011). The total market share of seeds with biotechnical traits is higher because the four biotechnology firms also license traits to independent seed firms. Our analysis documents how these forces are associated with how long a seed variety survives in the market.

which has been called “cannibalization effect” in the literature.

5 Seeds are an intermediate production input sold to farmers and so are different from final consumption goods such as beer and automobiles, the products analyzed in the other studies. However, in the context of survival, there are important similarities between consumers’ consumption decisions regarding final goods and farmers’ adoption decisions regarding seeds. Seed sales are based on final consumption through farmers’ derived demand. Consumption of durable goods such as cars also involves substantial R&D on the supply side, product marketing, and public information about performance, all of which likely influence purchase decisions and product survival.

6 The number of acres of U.S. corn planted with GM seeds rose from 4 percent in 1996 to 85 percent in 2009 (National Agricultural Statistics Service 2009).

7 They are Monsanto, Syngenta, DuPont, and Dow AgroScience. Another biotechnology company, Bayer CropScience, had entered the cotton seed market but had not (yet) entered the hybrid corn market during the study period.

8 As noted by an anonymous referee, the entry and exit of hybrids also may be affected by the quality of plant breeding programs that rely on classical genetics. However, except for increased use of GM technologies and structural changes associated with that use during the study period, we observed no major changes in how corn seeds were developed.
We use market and transaction-level data on corn hybrids that cover 2000 through 2007. Since most firms market more than one hybrid, we also examine the potential for intra-firm cannibalization effects. Our analysis suggests that product characteristics such as GM traits embedded in seeds and market characteristics such as competition and vertical structure significantly affect the survival of corn seed varieties. MKM found that expected lifespan for conventional varieties and GM hybrids were similar. In our study, which controls for the market's structure and spatial locations, we find that GM seeds tend to survive longer than conventional seeds in general. Moreover, market structure matters. Seeds supplied by a vertically integrated firm survive longer, and a firm's increased market share is associated with longer survival rates for its seeds. We also find evidence of information spillover effects in the success of GM seeds. We find no strong evidence of cannibalization effects but some indirect evidence is consistent with that hypothesis.

Note that the quantities and varieties of hybrid seeds available to farmers depend directly on trends in the total number of such hybrid varieties. The growing market share held by vertically integrated biotechnology-seed firms and longer survival rates associated with GM seeds suggest that, on a spatial level, independent regional seed firms may remain in the market for a while and that an adequate supply and number of varieties will be available in the near term but may not be for long. As the market share of these independent regional seed firms dwindles, many of them will not survive and farmers will have to rely on large firms that may or may not target spatially specific agronomic conditions in developing seeds. Our analysis suggests that such concerns are valid.

The United States Corn Seed Market

Our analysis relies on an extensive data set collected by Dmrkynetec (hereafter referred to as Dmrk) on the United States corn seed market. The data set comes from a stratified sample of U.S. corn farmers who were surveyed annually from 2000 through 2007 and provides farm-level information on corn seed purchases, acres of corn, hybrid identification numbers, biotechnical traits included in hybrids, and seed brands. It contains 168,862 observations on individual corn seed purchases from 48 states.

The U.S. hybrid corn seed market during the study period was characterized by a large number of hybrids. Since the initial commercialization of GM seeds in 1996, U.S. farmers have rapidly adopted this new technology. Figure 1, which is based on the Dmrk data, shows the number of conventional and GM corn hybrids purchased by U.S. farmers annually for 2001 through 2007. The number of hybrids remained around 3,600 between 2001 and 2004 and then increased substantially over the next three years. The number of GM hybrids was about 1,000 in 2001 and increased each year, rising to more than 4,000 in 2007. Thus, on average, more GM hybrids became available to farmers over time. This was good news for farmers who were in the process of adopting GM hybrids; on average, they had more options to choose from when deciding which hybrid(s) to plant. However, the fact that this was happening on average did not mean

---

9 Dmrkynetec was acquired by GfK Kynetec in May 2009 (www.gfk.com).
10 The survey is stratified to oversample large producers. The stratification scheme uses weights based on information from the U.S. Census of Agriculture (USDA).
that it was happening everywhere, leaving open the possibility of observing heterogeneous patterns across space and across hybrids. In particular, market structure could have positive or negative effects on the life cycle of specific hybrids.

Figure 2 plots the aggregate annual number of entering and exiting hybrids. A hybrid was defined as entering in year $i$ when it did not appear in the data set in any year prior to year $i$ and exiting in year $i$ when there was no purchase event in year $i$ and all subsequent years. There are more entering seeds than exiting seeds each year, and the overall trend for both events is increasing.

Since some of the hybrids observed in 2000 entered the market in prior years, there was a potential left-censoring bias in the data. Dmrk's data showed that more than 80 percent of the hybrids that exited had been in the market for three years or less. Thus, to mitigate the left-censoring problem, we estimate
the model using data for 2002 through 2007 and use the initial two years of data to establish hybrids’ entry points for calculating survival rates so that most of the seeds that entered the market prior to 2000 would have exited the market in 2000 and 2001. Our final data set contains 10,245 hybrids from 46 states—2,736 conventional seeds and 7,509 GM seeds; 5,081 hybrids exited during the data period and 5,164 remained in the market at the end of 2007.

Model and Discussion of the Data

Survival of a seed variety is determined by both demand and supply factors. On the supply side, costs, financing, and the firm’s R&D portfolio and marketing priorities and strategies may affect the product space that a firm decides to carry. Once a seed variety has been introduced, however, its survival in the market will depend on farmers’ adoption decisions. On the demand side, a variety’s survival can be influenced by consumers’ perceptions of and preferences for various attributes embedded in the seed, the availability of seeds in nonmarket channels (often related to the form of intellectual property protection associated with the seed technology), expected returns from planting the seed, and local agro-climatic conditions. Our empirical analysis uses market transaction data that represents equilibrium outcomes driven by both supply and demand factors. One limitation of this equilibrium data is that we cannot differentiate between the two groups of factors. Due to such limits on data availability, we rely on a reduced form specification in our survival analysis.

We begin with presentation of a standard Cox proportional hazard (CPH) model under an assumption of time-independent explanatory variables and confine our discussion to the case of corn hybrids. The lifespan of a corn hybrid variety, denoted by $t$, is defined as the number of years that the seed is observed in the market. We assume that $t$ is a realization of the random variable, $T$, following the probability distribution function $F(t) = P(T \leq t)$ and density function $f(t)$. The survival function is then defined as $S(t) = 1 - F(t)$, which measures the probability that a hybrid survives more than $t$ years. The hazard function considered at time $t$ is

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)},$$

which can be interpreted as the probability of exit given that the hybrid has already survived $t$ years. The CPH hazard function in logarithm form is specified as

$$\log(h(t, X_i)) = \log(h_0(t)) + \beta_i X_i$$

11 The MKM (2010) study examined the market for 1997 through 2009 and dropped the 1997 observations to avoid the left-censoring issue. We take a more cautious approach, using data from the first two years to identify entry events. The MKM study also dropped all observations of one-year survival on the basis that observation of a single year of survival could be due to survey error. For comparison, we re-estimated our model without the observations of a one-year survival time and found qualitatively similar results.

12 Structural modeling in an empirical survival analysis is uncommon primarily because the identification problems are empirically intractable. Many factors affect both supply and demand (e.g., a firm’s marketing strategies), and consumers’ adoption decisions may have a lagged feedback effect on which products firms choose to offer in subsequent periods.
where \( h_0(t) \) is the baseline hazard and \( X_i \) is a vector of time-independent explanatory variables.

Thus, for any two sets of predictors, \( X_i \neq X_i' \), the hazard ratio between the two observations in logarithm form is

\[
\log(HR) = \log\left(\frac{h_0(t) \exp(\beta_i X_i)}{h_0(t) \exp(\beta_i X_i')}\right) = \beta_i (X_i - X_i')
\]

where the hazard ratio (HR in equation 3) between the two sets of explanatory variables is constant over time. The hazard ratio is estimated by taking both exit and survival events and the corresponding survival time into account according to equation 1. We segregate the explanatory variables in \( X_i \) into three groups: (I) biotechnical characteristics of the seed, (II) market structure and information variables, and (III) location covariates.

GM technology creates substantial product differentiation in seeds by empowering various functions via individual biotechnical traits. Currently, there are two major groups of genes/traits in the GM seed market—insect resistance designed to reduce yield damage caused by insects and herbicide tolerance designed to reduce yield loss caused by competing plants (weeds)—that we incorporate into our model. For corn, insect resistance traits control damage caused by two insects, the European corn borer (represented in the model by variable \( ECB \)) and rootworm (represented in the model by variable \( RW \)). The herbicide tolerance traits (denoted as \( HT \)) work with the corresponding herbicides. With seeds with herbicide tolerance characteristics, farmers can apply the relevant herbicide to kill the weeds without damaging the crop.

GM seeds can contain a single trait or a “stack” of multiple traits from one or both of the two trait groups. In our model, the group I variables are associated with GM traits and consist of three dummy variables for GM traits that existed in the corn market during the study period (\( HT, ECB, \) and \( RW \)) and four additional dummy variables, one for each possible stacked combination: \( ECB_RW, ECB_HT, RW_HT, \) and \( ECB_RW_HT \). Thus, there are eight types of seed in our study: conventional seed (containing no GM traits) and the seven GM seed types.\(^{13}\)

Conventional seed is well represented in the data, and we use it as the base case for interpreting the marginal impacts of GM traits. When a corn hybrid contains one or more GM traits, the value of the corresponding dummy variable (\( HT, ECB, \) or \( RW \)) is one. If the seed is stacked with more than one trait, the value of the corresponding dummy variable (\( ECB_RW, ECB_HT, RW_HT, \) or \( ECB_RW_HT \)) is also one. Thus, for a single-trait \( ECB \) hybrid, \( ECB \) equals one and the values of the other seven dummy variables equal zero. For a stacked hybrid containing both insect resistance traits, \( ECB, RW, \) and \( ECB_RW \) equal one and the remaining variables equal zero. This structure allows us to obtain the marginal impact on survival of stacking both traits and determine how the combination of traits may add to or subtract from the marginal impact of the traits individually.

The variables in group II are associated with market structure and information. We include variables for the company’s state-level share of the market for the hybrid, a dummy variable identifying whether the firm is vertically integrated, the number of potential substitute hybrids produced by the same company, the

\(^{13}\) Our data set also contains observations of quadruple-stack seeds. We excluded them in this study because the quantity of those seeds was small and they came into the market mostly toward the end of our study period.
number of potential substitute hybrids produced by other companies, and the aggregate number of GM hybrids produced by all other companies. For each hybrid, these market variables take the value in the year in which the hybrid initially enters the market.\(^{14}\)

The market share variable is defined at a state level and is intended to capture the firm's market power position in that region. A firm's market dominance may affect its choice of marketing strategies and product spaces (e.g., Pinkse and Slade 2004), which in turn could affect survival of the seed variety. Firms often develop corn hybrids to perform well under specific agro-climatic conditions, and a hybrid developed for Wisconsin is unlikely to compete in the same market as a hybrid developed for California. We also include a dummy variable that represents whether the firm is vertically integrated with a biotechnology company. A vertically integrated firm may have an advantage over other firms associated with access to advanced technology and/or perceptions of farmers about such access, which may influence survival of the firm's lines of seed.

The remaining variables in group II provide a general framework for capturing changes in (i) the information set available to farmers, (ii) competition in the market, and (iii) the potential cannibalization effect, which may play a role in determining the lifespan of each seed product in the market. We define the close substitute hybrids as those hybrids in each of the eight possible seed categories previously mentioned. The number of close substitute hybrids produced by the same company and by other firms and the total number of GM seeds produced by other firms are incorporated solely as control variables. On one hand, close substitutes produced by a single firm could cannibalize the firm's previously released product lines, reducing their survival time (Asplund and Sandin 1999). On the other hand, a number of similar hybrids from one company could create positive information spillover among farmers, which could affect the firm's image and influence the survival rate of those seeds.\(^{15}\) Similarly, products offered by other firms can create competition pressure but also could provide information spillover among farmers adopting new GM technologies. Such information spillovers are often referred to as social learning or neighborhood effects in the literature on technology adoption (e.g., Foster and Rosenzweig 1995, Munshi 2004, Baerenklau 2005). All of these variables are constructed by counting the number of corresponding hybrids in the national market\(^{16}\) since farmers may gain perspective from information spillovers through various sources (e.g., trade shows, magazines, sales representatives, neighbors, and family).

GM traits can affect a farm's productivity in general by increasing its output and/or decreasing its use of inputs. However, no data were available on the performance of the corn hybrids analyzed in our study so the impacts of those quality characteristics are not estimated. Given the thousands of observations for each seed group in our data set, any effects of unobserved quality factors may be averaged out, in which case major differences in seed survival would

---

\(^{14}\) Since product exits can affect the value of these market variables, there is a potential endogeneity problem when using them to predict product exits. We use the value for the initial year only in this case to minimize endogeneity since it is unlikely that product exits in later years would affect the value of market variables in the year of entry.

\(^{15}\) For example, a firm that carries multiple hybrids in a given seed category could be viewed by farmers as having expertise in that particular technology.

\(^{16}\) We also estimated a model in which we included additional similar variables constructed at a state level. The results were qualitatively similar to the results presented here.
be attributable to embedded GM traits and institutional factors associated with horizontal and vertical structural changes in the seed market.

The variables in group III relate to geographic location. When unobserved seed performance parameters vary across space, we can control some of the effects of those parameters through latitude/longitude variables. We construct the weighted latitude and longitude for each hybrid using planted acres of it in different regions over time. For example, if a hybrid is observed in three locations (1, 2, 3) that each involve a unique number of acres, the weighted longitude is constructed as

$$\sum_{i=1}^{3} \text{longitude}_i \cdot s_i$$

where $s_i$ is the share of location $i$'s acres planted with the hybrid and

$$s_i = \frac{\text{acreage}_i}{\sum_{j=1}^{3} \text{acreage}_j}.$$  

Weighted latitudes are constructed similarly. Thus, the weighted variables reflect primary regions where the hybrid has been marketed. We allow for a nonlinear spatial pattern in the model by including both the linear and quadratic terms of the weighted latitude and longitude variables.17

The standard CPH model in equation 2 imposes an implicit assumption that all explanatory variables are time-independent such that the effects of those variables on the hazard are proportional over time. It follows that the hazard ratio in equation 3 is constant over time. The assumption of time-independence may not hold if a market is fundamentally changing in multiple ways. On the supply side, the market structure of the seed industry has changed dramatically during the study period, which may present a problem for the standard CPH model. However, the presence of new product features generally is not a problem provided that the underlying use of the product is not changing too much. Consider, for example, the retail market for laundry detergent. Laundry products are constantly upgraded to provide superior cleaning and/or stain-fighting abilities and claim to offer features such as better whitening and brightening agents or to work better in cold water. However, the market is fundamentally stable because products in the market space for cleaning clothes are all essentially designed to work well in automatic washing machines. In the case of corn hybrids, the emergence of genetic technologies has altered the way farmers think about producing corn but the hybrids are part of a production process that has not changed much during the years in question; corn must be grown in good soil and usually needs fertilizers and rain/irrigation to produce good yields. While GM traits offer different ways to manage pests and use herbicides, farmers are likely to continue to consider pest and weed management strategies in their production plans.

To test whether the assumption of time-independence holds for our explanatory variables, we perform a goodness-of-fit test based on Schoenfeld residuals (Schoenfeld 1982). The null hypothesis is that the proportional hazard holds over time. Three market structure variables fail the test at a

---

17 We also estimated the model with an additional interaction term of longitude and latitude. The results were robust and the interaction term was not statistically significant (p-value 0.87). For simplicity, we report the model without the interaction term.
5 percent significance level: firm market share, number of close substitute hybrids produced by the same firm, and number of substitute hybrids produced by other seed firms.

Kleinbaum and Klein (2005) identified three types of time-varying variables. Defined variables are those with time-varying effects by construction, internal variables come from a firm’s internal time-varying characteristics, and ancillary variables come from external time-varying characteristics. We suggest that a firm’s market share is a function of both internal and ancillary variables. Each firm’s internal characteristics and behavior likely impact its market share, as does the behavior of other firms. The number of intra-company close substitute seed lines is likely primarily an internal variable (although it may be affected indirectly by external market conditions), and the number of competitor hybrids is likely an ancillary variable as it may not be affected by own-firm behavior. Following Kleinbaum and Klein (2005), we interact each of the variables with the logarithm of survival time to manage time-varying effects. Equation 2 is modified into an extended CPH (ECPH) model:

\( \log(h(t, X_i, X_j(t))) = \log(h_0(t)) + \beta_iX_i + \beta_j(\log t \cdot X_j) \)

where \( X_j(t) \) contains the three covariates with time-varying effects and \( X_i \) includes all of the other covariates. The corresponding logarithmic hazard ratio function is

\( \log(HR) = \beta_i(X_i - X_i') + \beta_j(\log t \cdot (X_j - X_j')). \)

Our analysis of corn hybrid survival relies on equations 4 and 5. There are other survival models that can incorporate time-varying effects. However, the ECPH model was chosen for its robustness in estimating the regression coefficients and hazard ratios of interest given that the true parametric model of the seed industry may still be unclear because of the industry’s dramatic changes both structurally and technically in the past two decades.

Table 1 reports summary statistics of the variables used in the analysis. For seeds that entered the market on or after 2002 and exited the market on or before 2007, the mean survival time is 1.80 years for conventional hybrids and 1.63 years for GM hybrids. Of the seeds that had not exited by 2007, 79 percent are GM seeds. Thus, the survival rate for GM hybrids is likely biased downward since it only captures the lifetime of the early-exiting GM seeds and not the longer-lived GM seeds that exited after 2007.\(^{18}\) The average market share for the seed companies is 4.5 percent. Individually, market share varies substantially—from a dominant share of 91.1 percent in a local (state) market to a negligible share of close to zero.

On average, when entering the market, each hybrid faced 1,215 close substitute hybrids and 2,819 GM hybrids supplied by other companies while the firm supplied an average of 17 close substitutes nationally. On average, 9 percent of the number of hybrids in each local market were supplied by vertically integrated firms. The range of products provided by vertically integrated firms in a particular market (state) ranged from none to as much as

\(^{18}\) There were 5,081 hybrids in our sample that “failed”—that exited the market. At the end of 2007, 5,164 hybrids remained in the market (and thus were right-censored in our analysis), 4,096 GM hybrids and 1,068 conventional seeds.
Results

The ECPH model in equation 4 is estimated by maximizing the partial likelihood function\(^\text{19}\) based on the observed order of exits (Kleinbaum and Klein 2005). The estimated likelihoods and coefficients are presented in Table 2.\(^\text{20}\) Since the dependent variables in Table 2 are logarithms of hazard, they identify the degree of hazard with the hazard being the seed line’s exit from the market. Thus, a negative and significant coefficient implies that the degree of hazard is

\[^\text{19}\] The partial likelihood function can be specified as

\[
\log(L(\beta)) = \sum_i \beta X_i - \sum_{t \geq i} \exp(\beta X_i).
\]

We estimate the model using the Stata \textit{stcox} procedure, which controls for right-censoring and allows for time-varying variables.

\[^\text{20}\] We checked the collinearity diagnostic using the Stata \textit{collin} command. The average variance inflation factor was 1.53 (we excluded the square terms of latitude and longitude, which, by construction, were collinear with the linear terms of the two variables) so there was no evidence of collinearity in our independent variables. We also conducted a series of robustness checks, including Weibull, Compertz, exponential, clog-log, and logit models. Our results were robust to these checks.

Notes: There are eight hybrid groups and the number of seeds in each group varies substantially. Thus, the mean and standard deviations of the number of close substitute hybrids by other companies are pooled statistics of both cross-group and within-group variations.

Table 1. Summary Statistics of Selected Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Number of Observations</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hybrid survival (noncensored; in years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conventional</td>
<td>2,048</td>
<td>1.80</td>
<td>1.18</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>GM</td>
<td>4,963</td>
<td>1.63</td>
<td>1.02</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Market share (percent)</td>
<td>10,245</td>
<td>4.5</td>
<td>9.2</td>
<td>1.77E-05</td>
<td>91.1</td>
</tr>
<tr>
<td>Vertically integrated</td>
<td>10,245</td>
<td>0.09</td>
<td>0.28</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Number of close substitute hybrids by other companies</td>
<td>10,245</td>
<td>1,215</td>
<td>812</td>
<td>0</td>
<td>2,820</td>
</tr>
<tr>
<td>Number of close substitute hybrids by own company</td>
<td>10,245</td>
<td>17</td>
<td>24</td>
<td>1</td>
<td>193</td>
</tr>
<tr>
<td>Number of GM hybrids by other companies</td>
<td>10,245</td>
<td>2,819</td>
<td>1,028</td>
<td>1,302</td>
<td>4,206</td>
</tr>
<tr>
<td>Weighted latitude</td>
<td>10,245</td>
<td>41.53</td>
<td>2.47</td>
<td>27.60</td>
<td>48.77</td>
</tr>
<tr>
<td>Weighted longitude</td>
<td>10,245</td>
<td>91.23</td>
<td>5.85</td>
<td>70.53</td>
<td>124.75</td>
</tr>
</tbody>
</table>

90 percent of all seed. Finally, the weighted longitude and latitude information defined our study region as centered in Muscatine County, Iowa. The primary region in which the hybrids in our sample were grown ranged from Texas to North Dakota and from Maine to Washington.
## Table 2. Results from the Extended Cox Proportional Hazard Model

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>Specification 1 (S1)</th>
<th>Specification 2 (S2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>Z-statistic</td>
</tr>
</tbody>
</table>

### Group I: Biotechnical Characteristics – Benchmark Is Conventional Seeds

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Z-statistic</th>
<th>Coefficient</th>
<th>Z-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>GM</td>
<td>-0.471***</td>
<td>-5.80</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>ECB</td>
<td>—</td>
<td>—</td>
<td>-0.288***</td>
<td>-2.97</td>
</tr>
<tr>
<td>RW</td>
<td>—</td>
<td>—</td>
<td>-0.077</td>
<td>-0.56</td>
</tr>
<tr>
<td>HT</td>
<td>—</td>
<td>—</td>
<td>-0.391***</td>
<td>-4.56</td>
</tr>
<tr>
<td>ECB_RW</td>
<td>—</td>
<td>—</td>
<td>0.160</td>
<td>1.00</td>
</tr>
<tr>
<td>ECB_HT</td>
<td>—</td>
<td>—</td>
<td>0.374***</td>
<td>3.67</td>
</tr>
<tr>
<td>RW_HT</td>
<td>—</td>
<td>—</td>
<td>0.245*</td>
<td>1.61</td>
</tr>
<tr>
<td>ECB_RW_HT</td>
<td>—</td>
<td>—</td>
<td>-0.598**</td>
<td>-2.41</td>
</tr>
</tbody>
</table>

### Group II: Market Structure and Information Variables

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Z-statistic</th>
<th>Coefficient</th>
<th>Z-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market share – Share</td>
<td>-3.620***</td>
<td>-8.28</td>
<td>-3.571***</td>
<td>-8.11</td>
</tr>
<tr>
<td>Share *log(t)</td>
<td>2.175***</td>
<td>4.97</td>
<td>2.126***</td>
<td>4.83</td>
</tr>
<tr>
<td>Vertically integrated</td>
<td>-0.959***</td>
<td>-9.98</td>
<td>-0.875***</td>
<td>-8.70</td>
</tr>
<tr>
<td>Number of close substitute hybrids by own company – N_own</td>
<td>-0.01***</td>
<td>-7.15</td>
<td>-0.010***</td>
<td>-7.05</td>
</tr>
<tr>
<td>N_own*log(t)</td>
<td>0.003**</td>
<td>2.14</td>
<td>0.003**</td>
<td>2.17</td>
</tr>
<tr>
<td>Number of close substitute hybrids by other companies – N_others</td>
<td>-2.05E-04***</td>
<td>-4.74</td>
<td>-1.28E-04***</td>
<td>-2.63</td>
</tr>
<tr>
<td>Number of GM hybrids by other companies – N_GM_others</td>
<td>-3.16E-04***</td>
<td>-17.22</td>
<td>-2.76E-04***</td>
<td>-14.45</td>
</tr>
<tr>
<td>N_GM_others*log(t)</td>
<td>-3.93E-04***</td>
<td>-7.05</td>
<td>-4.22E-04***</td>
<td>-7.52</td>
</tr>
</tbody>
</table>

### Group III: Location Covariates

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Z-statistic</th>
<th>Coefficient</th>
<th>Z-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weighted latitude</td>
<td>-0.192**</td>
<td>-2.15</td>
<td>-0.217**</td>
<td>-2.41</td>
</tr>
<tr>
<td>Weighted latitude squared</td>
<td>0.002**</td>
<td>2.25</td>
<td>0.003**</td>
<td>2.51</td>
</tr>
<tr>
<td>Weighted longitude</td>
<td>-0.254***</td>
<td>-8.31</td>
<td>-0.259***</td>
<td>-8.43</td>
</tr>
<tr>
<td>Weighted longitude squared</td>
<td>0.001***</td>
<td>7.91</td>
<td>0.001***</td>
<td>8.03</td>
</tr>
</tbody>
</table>

Number of observations: 10,245 10,245
Log Likelihood: -43,707.86 -43,657.28

Notes: Statistical significance is noted by * at the 10 percent level, ** at the 5 percent level, and *** at the 1 percent level.
low relative to the benchmark of conventional hybrids and the GM hybrid will survive longer than the conventional seed. We consider two specifications for equation 4. In the first (S1), we use only a single dummy variable for GM seeds in group I, which allows us to investigate whether the survival rate of GM hybrids as a group is different from that of conventional hybrids. This specification applies a uniform survival rate to all of the GM characteristics. In the second specification (S2), we relax that restriction and allow the survival rate to vary by the seven GM characteristics in our model. Note that S2 involves interaction terms that link GM traits, time, and location of planting so the nominal values of some of the coefficients prevent us from making clear economic interpretations of their effects. To obtain meaningful results from these variables, we develop a simulation (described in the next section) that evaluates marginal changes in the hazard ratio presented in equation 5.

**GM Characteristics**

Conventional hybrids serve as the benchmark case by which we analyze dummy variables for the GM traits. The results from the first specification, shown in Table 2, suggest that GM seeds tend to survive longer than conventional hybrids; the coefficient on the GM dummy variable is negative and statistically significant. Moreover, the results vary by GM trait type. In the second specification, the estimated coefficients on two of the individual GM characteristics, ECB and HT, are negative and statistically significant, indicating that those single-trait seeds survive longer in the market than conventional hybrids. The single-trait RW seed’s survival rate is not different from conventional seeds. The coefficient on ECB-HT is positive and significant while the coefficient on ECB_RW_HT is negative and significant. We analyze the full marginal effects of the stacked-trait seeds in the next section.

**Market Structure and Information**

The five market-level variables and additional interaction terms that capture the time-varying nature of market share, number of own-firm substitutes, and number of other-firm substitutes are presented in Table 2. Market share, (market share)*log(t), and vertical integration allow us to evaluate hypotheses about market structure. For these variables, the results from the two specifications are very similar. The coefficient of market share is negative and statistically significant, a result that is consistent with Greenstein and Wade (1998). A greater market share is associated with a lower degree of hazard. However, the coefficient of the time interaction term is positive and significant, suggesting that this effect decreased over time. The results also indicate that a seed company’s vertical integration with a biotechnology company reduces the hazard of the company’s products exiting the market. Anecdotal evidence suggests that vertically integrated seed companies have access to higher quality traits and seed germplasms than seed companies operating under a license from a biotechnology company (Chataway and Tait 2000). Shi and Chavas (2011) examined the market for soybean seeds and Shi,
Stiegert, and Chavas (2011) analyzed the market for cotton seeds. Both found that seed prices tended to be higher under vertical integration than under licensing. If higher prices reflect better quality associated with seeds supplied via vertically integrated channels, such seeds may survive longer in the market as long as they offer a greater net benefit to farmers than lower quality, less expensive seeds do.

Three variables and two associated interaction terms evaluate the impact of information and product cannibalization. The coefficient for the number of own-firm close substitutes \((N_{\text{own}})\) in the market was negative and significant while the coefficient for the time-variant interaction term \((N_{\text{own}} \times \log(t))\) was positive and significant. Since the number of close substitutes by the firm increased over time, survival strengthened, suggesting that farmers tended to purchase seeds from firms that enlarged the scope of their offerings. Such purchases may be a result of a growing positive reputation for the firm or increased information about product superiority. The positive and significant time-varying counterpart to \(N_{\text{own}}\) provides some limited support for a cannibalization effect. It suggests that the positive impact of a firm’s close substitutes on survival decreases over time, which may be associated with internal cannibalization.

The coefficient on the number of close substitutes marketed by competing firms \((N_{\text{others}})\) is negative and statistically significant. Thus, the survival time of seed products increased with an increasing number of close substitutes. The coefficients on the number of GM hybrids produced by other firms \((N_{\text{GM others}})\) and its time-varying interaction term are both negative and significant. As the number of GM hybrids in the market increased, the rate of survival of GM hybrids also increased and this effect grew over time. If we view GM seeds as imperfect substitutes and combine them with close substitutes for a given GM seed, our results for \(N_{\text{others}}\) and \(N_{\text{GM others}}\) contradict our hypothesis about the effect of competition—that increased competition leads to faster product turnover. Alternatively, perhaps effects of positive learning spillover regarding GM technologies led to longer survival rates. It seems that farmers pay attention to trends in adoption of GM seeds and relate their own purchases to general acceptance of new seed technologies.

**Location Covariates**

We controlled for spatial effects using linear and quadratic terms of the longitude and latitude of primary planting regions. Again, the results from the two specifications are very similar. Given the weighting scheme we employed for where the seeds were used, the results provide a general guide to where survival rates were longer or shorter than the baseline. Since the early development of hybrid technology in the 1930s, firms have developed and marketed new hybrids to meet the specific needs of farmers in various regions (Griliches 1960). The introduction of GM technology did not change the region-specific nature of seed development. In our model, the coefficients of the linear terms are negative while the quadratic terms are positive and all of the coefficients are significant at the 5 percent level. Geographically, hazard of exit decreases from east to west and from south to north. It peaks in the southeasterly portion of the Corn Belt and then declines to the west and north. These results suggest that corn hybrids survive longest in the southeastern region of the Corn Belt. Overall, seeds developed for and marketed in outlying...
production regions such as California and North Dakota do not survive as long as those developed for the Corn Belt.

Implications

To quantify the marginal effects of the main explanatory variables, we simulate how changes in the variables in groups I and II would induce changes in the hazard ratio. We set the benchmark case as a conventional hybrid introduced by a seed company that is not vertically integrated with a biotechnology company and hold all other continuously valued explanatory variables in groups II and III at their means. We calculate the hazard ratio, 

\[ HR = \frac{h(t, \tilde{X}, \tilde{X}_i(t))}{h(t, X, X_i(t))} \]

by adjusting the relevant dummy variable from zero to one or by increasing the relevant continuous variable by one standard deviation from its mean. The results are reported in Table 3.

Table 3. Marginal Effects on the Hazard Ratio

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Benchmark Scenario</th>
<th>New Scenario</th>
<th>Hazard Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group I: GM Characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECB</td>
<td>-0.288***</td>
<td>Conventional</td>
<td>Single ECB</td>
<td>0.75***</td>
</tr>
<tr>
<td>RW</td>
<td>-0.077</td>
<td>Conventional</td>
<td>Single RW</td>
<td>0.93</td>
</tr>
<tr>
<td>HT</td>
<td>-0.391***</td>
<td>Conventional</td>
<td>Single HT</td>
<td>0.66***</td>
</tr>
<tr>
<td>ECB_RW</td>
<td>0.160</td>
<td>Conventional</td>
<td>ECB_RW</td>
<td>0.81</td>
</tr>
<tr>
<td>ECB_HT</td>
<td>0.374***</td>
<td>Conventional</td>
<td>ECB_HT</td>
<td>0.74***</td>
</tr>
<tr>
<td>RW_HT</td>
<td>0.245*</td>
<td>Conventional</td>
<td>RW_HT</td>
<td>0.80</td>
</tr>
<tr>
<td>ECB_RW_HT</td>
<td>-0.598**</td>
<td>Conventional</td>
<td>ECB_RW_HT</td>
<td>0.26***</td>
</tr>
<tr>
<td>Group II: Market Structure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share</td>
<td>-3.571***</td>
<td>4.5 percent</td>
<td>13.7 percent</td>
<td>0.82***</td>
</tr>
<tr>
<td>Share*log(t)</td>
<td>2.126***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N_own</td>
<td>-0.010***</td>
<td>17</td>
<td>41</td>
<td>0.83***</td>
</tr>
<tr>
<td>N_own*log(t)</td>
<td>0.003**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N_GM_others</td>
<td>-2.76E-04***</td>
<td>2,819</td>
<td>3,847</td>
<td>0.56***</td>
</tr>
<tr>
<td>N_GM_others*log(t)</td>
<td>-4.22E-04***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vertically integrated</td>
<td>-0.875***</td>
<td>Not integrated</td>
<td>Integrated</td>
<td>0.42***</td>
</tr>
</tbody>
</table>

Notes: Statistical significance is noted by * at the 10 percent level, ** at the 5 percent level, and *** at the 1 percent level.

22 If the variable is time-dependent, we set the time index at \( t = 2 \).
The first seven rows in Table 3 report the marginal effects of GM characteristics on the hazard ratio of seed survival. All of the GM seeds have a hazard ratio of less than one, which suggests that GM seeds survive longer in the market than conventionally bred hybrids. The hazard ratio is statistically significant for only one of the hybrids that contain the rootworm resistance trait (\textit{RW})—the triple-stack seed (\textit{ECB_RW_HT}, 0.26). This result suggests that the hazard of market exit for seeds with the rootworm trait is the same as the hazard for conventional seeds. Anecdotal evidence also has suggested that the rootworm trait has not had as much success in the market as the trait for resistance to the corn borer. The hazard ratio of seed with the \textit{ECB} trait only is 0.75 and the ratio of seed with the \textit{HT} trait only is 0.68. Thus, inserting the single \textit{ECB} or \textit{HT} trait into a conventional hybrid would increase the survival rate of the resulting GM seed by 25 percent and 32 percent respectively, \textit{ceteris paribus} (i.e., $1.00 - 0.75 = 0.25; 1.00 - 0.68 = 0.32$).

Embedded GM traits increase product differentiation in the seed market, which may in turn increase the survival of GM seeds relative to conventional seeds.\textsuperscript{23} However, we find that double stacking GM traits does not extend survival rates (it does not reduce survival rates either). The hazard ratios of double-trait seeds that included the \textit{RW} trait (\textit{ECB_RW} and \textit{RW_HT}) are less than one but not significant. \textit{ECB_HT} seeds survive longer than conventional seeds but the survival rate is not much different than the rate for single-trait \textit{ECB} and \textit{HT} seeds. This lack of increase in survival time may be related to reductions in yield that tend to occur in multi-trait GM seeds. Or there might be agronomic limits on survival that cannot be extended much past what is observed in single-trait seeds.

The triple-trait seed generates the lowest hazard ratio (0.26), implying that its rate of survival is 74 percent greater than the rate for the conventional hybrid. This seed type did not emerge until later in the period covered by the data set; therefore, we take a cautionary view of this finding. It could mean that farmers were experimenting with triple-trait technology and that their interest in it would change after they had more experience with its cost/benefit tradeoffs.

We turn next to the market variables in group II. When the market share of a seed company during the hybrid’s entry year exceeds the mean of 4.5 percent by one standard deviation (rising to at least 13.7 percent), the rate of survival of that hybrid increases by 18 percent, \textit{ceteris paribus}. Thus, an expanding market position may help a firm extend the life of its products in the market. Additionally, seeds from vertically integrated firms survive longer than seeds from other firms by 58 percent. Major players in the market such as Monsanto, Syngenta, DuPont, and Dow AgroScience are vertically integrated and thus may have a competitive advantage over smaller regional seed companies. If vertically integrated firms’ products remain in the market longer, those firms may be able to recover their R&D costs via lower prices than they could otherwise or may enjoy increased returns to R&D, which could drive independent regional

\textsuperscript{23} Conventional seeds are often protected by plant variety protection (PVP), a form of patent that allows farmers to use saved seeds and preserves breeders’ rights for research purposes. GM seeds are often protected by a utility patent that basically rules out the two exemptions provided by PVP and is a much stronger protection than PVP. This difference in intellectual property protection mechanisms may affect how firms appropriate economic returns and will affect the survival of hybrids in the market. This difference is not as prominent for corn as it is for other GM crops such as cotton and soybeans. The hybrid technology in conventional corn seed provides natural protection that prevents farmers from using saved seed—the yields from saved seed are much smaller and considerably less predictable than yields from first-generation hybrid seed.
firms out of the market. Figures 1 and 2 suggest that independent regional seed firms can, in the near term, continue to survive in the seed market and that availability of hybrids to farmers will remain adequate. However, independent firms may not be able to survive in the long run.

We found considerable evidence of a strong spillover effect in the remaining variables in group II. Each variable provides information about the general acceptability of close substitutes for a hybrid and of GM technology. An increase of one standard deviation in the number of close substitutes within a company increases the survival of the hybrid by 17 percent. If we do not consider the counteracting time effects, the hazard reduction would be 22 percent.24 When the number of close substitutes produced by other companies is increased by one deviation, the survival rate for the hybrid increases by 10 percent. Finally, the number of GM hybrids produced by other companies has a strong effect on the hazard of a product exiting the market. An increase of one standard deviation generates a 44 percent increase in the survival rate. Thus, it appears that the survival rate of a GM hybrid depends on a general sense among farmers that GM seeds are growing in acceptance throughout the United States and on the firm’s reputation, market share, and backward vertical integration with a biotechnology firm.

Conclusion

In just the past decade, product offerings in the U.S. corn hybrid seed industry have shifted dramatically toward GM hybrids. We investigated determinants of survival of corn hybrids in the U.S. seed industry using transaction-level data for 2000 through 2007. We find that these new seed characteristics—the embedded GM traits—have contributed to the survival of corn hybrids in general relative to conventional seed by differentiating products in the market, primarily via the advent of single-trait GM seeds. However, the development of multi-trait hybrids did not result in an extension of survivability for those products during the study period. In addition, products from firms with larger market shares and from firms that are vertically integrated with a biotechnology company tend to have longer product lifespans than products from smaller firms that are not vertically integrated. These results related to firm-level characteristics imply that longer survival rates could lead to significant cost advantages for larger vertically integrated firms that could subsequently drive smaller independent seed firms out of business. And the resulting concentration of suppliers could eventually mean fewer choices for farmers and/or higher prices.

There is evidence of positive spillover learning effects. Survival rates increase when a firm offers a larger number of choices, when more close substitutes are available, and when the number of GM seeds in the market increases.

Our results shed light on factors that determine the rate of survival of seed products in the U.S. corn hybrid market. Farmers have benefitted greatly from a vibrant private hybrid seed industry. Extensive competition among firms has pressured seed companies to continually develop better products that offer higher yields for growers in all of the major growing regions of the country. Concern has been expressed about whether the increasing market share of vertically integrated seed/biotechnology firms will stifle that development

24 The counteracting time effect increases over time. At $t = 5$ (the maximum survival time we can observe in the model), the hazard reduction would be 10 percent.
pressure, resulting in excessively long survival rates, fewer seed choices as smaller independent regional seed firms fail, inflated prices, and less R&D devoted to increasing corn yields. Our results suggest that such concerns are valid.

Our analysis can be extended by collecting more comprehensive data on seed performance. Specifically, it would be useful to include spatially oriented data on the performance of hybrids in testing prior to their release for sale and the location-specific performance of previously released hybrids that includes rotations, soil quality, irrigation, and herbicide and pesticide applications. Also helpful would be a separate exploration of how supply-side and demand-side factors affect seed survival patterns. As the corn seed market continues to evolve, research on how large biotechnology firms influence the product space for hybrid corn seeds in the market will be important as well.

References


