R&D Concentration under Endogenous Fixed Costs: Evidence from the Agricultural Biotechnology Industry

Benjamin C. Anderson
Colgate University, bcanderson@colgate.edu

Ian M. Sheldon
The Ohio State University, sheldon.1@osu.edu

October 2013

Abstract

We derive the theoretical lower bound to R&D concentration and empirically testable hypotheses for an industry characterized by endogenous fixed costs (EFC). Using data on field trial applications of genetically modified (GM) crops, we estimate the lower bound to R&D concentration in the agricultural biotechnology sector. The empirical results imply that the markets for GM corn, cotton, and soybean seeds are characterized by EFC with predicted lower bounds to the one-firm R&D concentration ratio for an infinitely-sized market of 54.8%, 47.3%, and 78.6%, respectively. We find no evidence that adjusting for merger and acquisition activity significantly increases the lower bounds.

Keywords: *market structure, R&D, endogenous fixed costs, agricultural biotechnology*

JEL Codes: L22, Q16

Acknowledgements:
This project was supported by the Agricultural Food Research Initiative of the National Institute of Food and Agriculture, USDA, Grant #2008-35400-18704. Previous versions of this article have been circulated with the title "Endogenous R&D Investment and Market Structure: A Case Study of the Agricultural Biotechnology Sector". The authors wish to thank participants at the 9th Annual International Industrial Organization Conference, April 8-10, 2011, Boston, MA and at the 82nd Annual Meetings of the Southern Economic Association, November 16-18, 2012, New Orleans, LA.

I. Introduction

In 2009, the US Department of Justice and US Department of Agriculture announced a series of joint public workshops exploring issues of competition in agriculture. (DOJ Press Release, 2009) The public comments submitted in support of these workshops raised concerns of anticompetitive practices and "excessively" high levels of concentration across a variety of agriculture-related industries with many focusing specifically on the seed industry and agricultural biotechnology firms. A group of fourteen state attorney generals were one such group raising concerns of concentration in the seed industry citing that "increased vertical integration and acquisitions may have raised the bar for entry so high that entry into the *trait market* is difficult, or nearly impossible" [emphasis ours]. (Munson, London, and Lindeback, 2010) The comments submitted by the Farmer to Farmer Campaign on Genetic Engineering, a "network of 34 farm organizations…that (*sic*) seeks to build a farmer driven campaign focused on concerns around agricultural biotechnology", encapsulate many of the public concerns of concentration in the seed industry. (Hubbard, 2009) However, as the report identifies, one of the primary concerns regarding concentration in the seed industry is not solely the levels of market concentration as measured by sales, but rather the concentration in seed traits and germplasm.

Over the past three decades, agricultural biotechnology, a sector characterized by rapid innovation and firm consolidation, has become an increasingly important component in agricultural production. Gaisford, et al. (2001) define (agricultural) biotechnology as "the use of information on genetically controlled traits, combined with the technical ability to alter the expression of those traits, to provide enhanced biological organisms, which allow mankind to lessen the constraints imposed by the natural environment." Our analysis focuses on agricultural biotechnology firms that develop genetically-modified (GM) corn, cotton, and soybean seed varieties for commercialization purposes. In the US in 2006, private-sector firms spent $1.2 billion in crop seed research and development (R&D) which accounts for half of all R&D expenditure on crops and one-fifth of all R&D expenditure on food and agriculture. (Fuglie, et al., 2011)

In the agricultural biotechnology crop seed sector, the four-firm market concentration ratio increased from 21.1% in 1994 to 53.9% in 2009. (Fuglie, et al., 2011) However, increased concentration has had an ambiguous effect on R&D investment as R&D intensity, measured as the share of industry-level R&D expenditure to sales, increased from 11.0% in 1994 to 15.0% in 2000 before falling back to 10.5% in 2009. (Ibid) Moreover, these aggregate numbers obscure significant heterogeneity across firms as they vary by size and innovation strategy with the eight largest seed companies have an average R&D intensity of 15.8% while accounting for 75.6% of the global R&D share. (Ibid) Figure 1 plots the three-year average one-firm concentration ratios and adoption rates for GM crops for each crop type from 1996-2010. The graph illustrates two important trends: (i) increasing rates of adoption of GM seed varieties across time; and (ii) single firm R&D concentration ratios that initially increased and have remained consistently high across time.

Using data on field trial applications for GM crops, we exploit variations in technology and market size across time and sub-markets to analyze whether the agricultural biotechnology sector is characterized by a lower bound to concentration consistent with an endogenous fixed cost (EFC) framework. If the agricultural biotechnology sector is characterized by endogenous fixed costs, the high levels of concentration, accompanied with high levels of innovative activity amongst the largest firms, are a natural outcome of technology competition. Therefore, as market size becomes large, these industries are characterized by firms that are able to gain an increased market share by investing in product quality and preclude entry by additional competitors, thus bounding concentration away from perfectly competitive levels.

The model of endogenous market structure and R&D investment developed by Sutton (1998) predicts a lower bound to firm R&D intensity that is theoretically equivalent to the lower bound to firm concentration under significantly large markets. We draw upon the results of Sutton (1998) in order to determine the empirical predictions of the EFC model regarding R&D concentration, defined as firm R&D relative to industry R&D. The empirical predictions imply that: (i) the lower bound to R&D concentration is convergent in market size (i.e., the theoretical lower bound is not independent of the size of the market as is the case with sales concentration); and (ii) R&D concentration moves in an opposite direction from firm concentration with changes in market size such that larger markets are characterized by greater concentration in R&D.

We use data on R&D investments, in the form of field trial applications for genetically modified (GM) crops, to test for lower bounds to R&D concentration among agricultural biotechnology firms. Using cluster analysis, we define regional sub-markets for each GM crop type (corn, cotton, and soybeans) based upon observable data on farm characteristics and crop production practices at the state level. We exploit variation along two dimensions: (i) geographically as adoption rates for GM crop varieties vary by state and agricultural region; and (ii) inter-temporally as adoption rates for GM crops have been steadily increasing over time. We estimate the lower bounds to R&D concentration using a two-step procedure suggested by Smith (1994) in order to test whether the single firm

R&D concentration ratios follow an extreme value distribution. Additionally, we examine the specific concern that mergers and acquisitions have increased the concentration of innovative activity and intellectual property in the agricultural biotechnology sector, an issue discussed in Moss (2009), Dillon and Hubbard (2010), and Moschini (2010).

The results from the empirical estimations support the hypothesis that the GM corn, cotton, and soybean seed markets are characterized by endogenous fixed costs to R&D with the theoretical lower bounds to R&D concentration ranging from 54.8% for corn, 47.3% for cotton, and 78.6% for soybeans. In general, our results are robust to alternate definitions of the data or model specifications. Finally, we find little to no evidence that accounting for mergers and acquisitions significantly increases the lower bound to R&D concentration. Overall, the results reveal the importance of sunk, fixed R&D investments in jointly determining both the levels of concentration and innovation activity.

To our knowledge, ours is the first examination of a specific industry in the context of firm-level investments in R&D, although the empirical analysis of Marin and Siotis (2007) of chemical manufacturers does differentiate between product markets characterized by high and low R&D intensities. We contribute to the industrial organization literature by applying an EFC model to a previously unexamined industry as well as derive and estimate the lower bound to R&D concentration under endogenous fixed costs. Additionally, our analysis extends previous examinations into merger and acquisition activity in agricultural biotechnology in estimating whether firm consolidation has had a significant impact upon the observed patterns of R&D concentration while abstaining from addressing the possible causal mechanisms behind the consolidation activity. Our results are of interest to both regulators and policymakers concerned with the observed high levels of concentration, and in particular the high levels of concentration in intellectual property, in agricultural biotechnology.

The remainder of the analysis is organized as follows: the second section presents a brief overview of agricultural biotechnology and discusses the literature addressing innovation and concentration in GM crops; the third section develops the theoretical model of R&D concentration and derives the empirically testable hypotheses; the fourth section discusses field trial applications as well as the other data utilized in the estimation; the fifth section presents the results and discussions; and the final section concludes.

II. What is Agricultural Biotechnology?

Prior to the 1970s, the development of new plant varieties was largely limited to Mendelian-type genetics involving selective breeding within crop types and hybridization of characteristics to produce the desired traits. Generally, it was impossible to observe whether the crops successfully displayed the selected traits until they had reached maturity implying a considerable time investment with each successive round of experimentation. If successful, additional rounds of selective breeding were often required in order to ensure that the desired characteristics would be stably expressed in subsequent generations. This process is inherently uncertain as crop scientists and breeders rely upon "hit-and-miss" experimentation, implying that achieving the desired outcome might require a not insubstantial amount of time and resources.

The expansion of cellular and molecular biology throughout the 1960s and 1970s, specifically the transplantation of genes between organisms by Cohen and Boyer in 1973, increased the ability of crop scientists to identify and isolate desired traits, modify the relevant genes, and to incorporate these traits into new crop varieties via transplantation with greater precision (ISAAA, 2010). These advances, along with the U.S. Supreme Court decisions in *Diamond v. Chakrabarty* (1980) and *J.E.M. Ag Supply v. Pioneer Hi-Bred* (2001) regarding the patenting of genetically engineered organisms, had two key implications for agricultural seed manufacturers and plant and animal scientists (Moschini, 2010). First, the ability to identify and isolate the relevant genetic traits greatly facilitated the transference of desirable characteristics through selective breeding. Second, the ability to incorporate genetic material from one species into the DNA of another organism allowed for previously infeasible or inconceivable transfers of specific traits. Perhaps the most widely known example of this was the incorporation of a gene from the soil bacterium *Bacillus thuringiensis* (Bt) that produces the Bt toxin protein. This toxin is poisonous to a fraction of insects, including the corn borer, and acts as a “natural” insecticide. When the gene is incorporated into a plant variety, such as corn, cotton, and now soybeans, the plants are able to produce their own insecticides, thereby reducing the need for additional application of chemical insecticides.

Rapid technological innovation and observed firm consolidation has led to several empirical examinations of market structure and innovation in the agricultural biotechnology industry. The industry attributes consistently identified by the literature and that factor into the proposed analysis include: (i) endogenous sunk costs in the form of expenditures on R&D that may create economies of scale and scope within firms[1]; (ii) seed and agricultural chemical technologies that potentially act as complements within firms

[1] Chen, Naseem, and Pray (2004) find evidence that supports economies of scope as well as internal and external spillover effects in R&D. However, they fail to find economies of scale or a correlation between R&D and the size of agricultural biotechnology firms.

and substitutes across firms; (iii) property rights governing plant and seed varieties that have become more clearly defined since the 1970s; and (iv) high levels of consolidation activity in the form of mergers and acquisitions. We extend the stylized facts for the agricultural biotechnology industry by identifying the relevance of sunk costs investments in R&D in the joint determination of market concentration and innovation.

Among the literature that examines the relationship between market structure and innovation, Schimmelpfennig, Pray, and Brennan (2004) find a negative and endogenous relationship between measures of industry concentration and R&D intensity whereas Brennan, Pray, and Courtmanche (1999) find a heterogeneous impact upon innovation activity following mergers and acquisitions. The explanations behind the high levels of firm consolidation activity in the agricultural biotechnology industry have included the role of patent rights in biotechnology (Marco and Rausser, 2008), complementarities in intellectual property in biotechnology (Graff, Rausser, and Small, 2003; Goodhue, Rausser, Scotchmer, and Simon, 2002), and strategic interactions between firms (Johnson and Melkonyan, 2003). Additional stylized examinations have identified an endogenous, cyclical relationship between industry concentration and R&D intensity (Oehmke, Wolf, and Raper, 2005), decreasing product life cycles associated with increasing innovation in corn seed (Magnier, Kalaitzandonakes, and Miller, 2010), and an endogenous relationship between firm innovation strategies, including complementary intellectual assets, and industry consolidation characteristics (Kalaitzandonakes and Bjornson, 1997). Whereas previous examinations have focused upon identifying the endogenous relationship between R&D intensity and concentration in agricultural biotechnology, we determine whether (sunk) R&D investments drive this relationship.

III. Endogenous Market Structure and Innovation: The "Bounds" Approach

A Lower Bound to R&D Concentration

We adapt the theoretical endogenous fixed cost model of market structure and sunk R&D investments developed by Sutton (1998) to derive the empirically testable hypotheses for the lower bound to R&D concentration under endogenous fixed costs. Sutton (1998) finds that the lower bound to market concentration, defined as the ratio of firm sales to total industry sales, is bounded away from perfectly competitive levels when an industry is characterized by endogenous fixed costs (EFC). Moreover, the EFC theory implies that R&D intensity, defined as the ratio of firm R&D to firm sales, is also characterized by a lower bound which is equivalent to the bound to market concentration as markets become large. However, the theory fails to address the implications of the EFC model on R&D concentration, defined as the ratio of firm R&D to industry R&D, within these industries.

The specification of the theoretical model, and subsequent empirical analysis, relies upon a set of assumptions regarding the nature of product differentiation in the agricultural biotechnology sector. First, we assume that there exist regional variations in the demand for specific seed traits, such as herbicide tolerance or insecticide resistance, and that these regional variations create geographically distinct sub-markets. This assumption corresponds with the empirical findings of Stiegert, Shi, and Chavas (2011) and Shi, Chavas, and Stiegert (2010) of spatial price differentiation in GM corn and implies that the agricultural biotechnology industry is characterized by horizontal product differentiation. Additionally, Shi and Chavas (2011) and Shi, Stiegert, and Chavas (2009) allow for seed prices to vary by regional market structure characteristics for soybean seeds and cotton seeds, respectively. Secondly, we assume that farmers value higher quality products such that a firm competes within each sub-market primarily via vertically differentiating the quality of its seed traits. Thus, we estimate a model of vertical product differentiation in the agricultural biotechnology sector while accounting for horizontal differentiation via the definition of geographically distinct product sub-markets.

For simplicity of analysis, it will be beneficial to introduce notation for industry sales revenue and R&D expenditure. Following the notation for the sales revenue Π_{im} and R&D expenditure F_{im} for some firm i in sub-market m, we define total industry sales revenue Π_m and R&D expenditure F_m in sub-market m by summing across all firms such that $\Pi_m = \sum_{i \in m} \Pi_{im}$ and $F_m = \sum_{i \in m} F_{im}$. Additionally, we define the degree of market segmentation (or product heterogeneity) $h_m \in [0,1]$ as the share of industry sales revenue in sub-market m accounted for by the largest product category l such that:

$$h_m = \max_l \frac{\Pi_{lm}}{\Pi_m}, \tag{1}$$

where $h_m = 1$ corresponds with a sub-market in which only a single product is offered.

Drawing upon the non-convergence results (Theorems 3.1-3.5) of Sutton (1998), the lower bound to the single firm concentration ratio C_{1m} for the quality-leading firm in sub-market m can be stated as:

$$C_{1m} = \frac{\widehat{\Pi}_m}{\Pi_m} \geq \alpha(\sigma, \beta) \cdot h_m, \tag{2}$$

where α is some constant for a given set of parameter values (σ, β), σ is a parameter capturing consumer preferences and product market substitutability, β is a parameter capturing the elasticity of R&D expenditures, and hat accent characters $(\widehat{\ })$ correspond to the variables of the quality-leading firm. Thus, the value of alpha α depends upon industry technology, price competition, and consumer preferences and signifies the extent that a firm can escalate quality via R&D investment and capture greater market share from rivals. Equation (2) implies that the lower bound to market concentration is independent of the size of the market in endogenous fixed cost industries, whereas under exogenous fixed cost industries, the lower bound to market concentration is decreasing and approaches zero as the size of the market increases.

Moreover, Sutton's (1998) Theorem 3.2 implies an equivalent expression for the lower bound to R&D-intensity $\hat{\mathrm{P}}_m$ for the quality-leading firm such that:

$$\hat{\mathrm{P}}_m = \frac{\hat{F}_m}{\hat{\Pi}_m} \geq \alpha(\sigma,\beta) \cdot h_m - \frac{F_0}{\Pi_m}. \tag{3}$$

Equation (3) implies that the R&D/sales ratio shares the same lower bound as the single firm concentration ratio as the size of the market becomes large (i.e., $\Pi_m \to \infty$).

The lower bound to R&D concentration can be derived directly by multiplying both sides of equation (3) by the sales revenue of the market leading firm in sub-market m:

$$\hat{F}_m \geq \alpha(\sigma,\beta) \cdot h_m \cdot \hat{\Pi}_m - F_0 \cdot \frac{\hat{\Pi}_m}{\Pi_m}. \tag{4}$$

Dividing both sides of equation (4) by total industry sales revenue in sub-market m yields:

$$\frac{\hat{F}_m}{\Pi_m} \geq \left[\alpha(\sigma,\beta) \cdot h_m - \frac{F_0}{\Pi_m}\right] \cdot \frac{\hat{\Pi}_m}{\Pi_m}. \tag{5}$$

However, free entry in equilibrium implies that total industry sales revenue Π_m equals total industry R&D expenditure F_m such that equation (5) can be written as:

$$\frac{\hat{F}_m}{F_m} \geq \left[\alpha(\sigma,\beta) \cdot h_m - \frac{F_0}{\Pi_m}\right] \cdot \frac{\hat{\Pi}_m}{\Pi_m}. \tag{6}$$

Defining the ratio of R&D concentration for the quality-leading firm as $R_{1m} = \frac{\hat{F}_m}{F_m}$ and substituting for equation (2), the condition on the lower bound to the single-firm concentration ratio is:

$$R_{1m} \geq \left[\alpha^2(\sigma,\beta)h_m^2 - \alpha(\sigma,\beta)h_m \frac{F_0}{\Pi_m}\right]. \tag{7}$$

Equation (7) provides the empirically testable hypothesis for endogenous fixed costs relating the lower bound to concentration in R&D expenditure to market size, the minimum R&D setup cost, and the level of product heterogeneity. If sunk R&D costs are endogenous, there would be a nonlinear relationship between the degree of market segmentation (product homogeneity) h_m and the concentration of R&D R_{1m} for a given market. Moreover, equation (7) implies a lower bound to the ratio of R&D concentration that converges to some constant $\alpha^2(\sigma,\beta)h_m^2$ as the size of the market becomes large. For finitely sized markets though, the lower bound to R&D concentration is increasing in market size such that R&D expenditures are less concentrated in smaller sized markets.

If the industry is instead characterized by exogenous fixed costs, then the ratio of R&D concentration in sub-market m can be expressed as:

$$R_{1m} = \frac{\hat{F}_m}{F_m} = \frac{F_0}{\Pi_m}. \tag{8}$$

For some minimum fixed setup cost F_0, concentration in R&D investments is decreasing in market size and approaches 0 as market size becomes large and, contrary to the case of endogenous fixed costs, the R&D concentration under exogenous fixed costs is greatest in small markets.

Figure 2 compares the lower bounds to R&D concentration for industries characterized by low and high levels of product heterogeneity h for a range of α

parameters as market size Π increases. If an industry is characterized by homogenous products (i.e., low $h = 0.75$), there is no range of α such that firms invest more in R&D in excess of the minimum setup cost associated with entry. However, if an industry is characterized by differentiated products (i.e., high $h = 0.25$) and sufficiently large α, then there is an incentive for firms to escalate R&D investment to increase product quality such that R&D concentration remains bounded away from zero as market size increases.

Empirical Specification

The empirical specification that we adopt was developed in Sutton (1991) and has since been adapted and extended in Giorgetti (2003), Dick (2007), and Ellickson (2007). This framework has been utilized to empirically examine a variety of industries including food and beverage manufacturers (Sutton, 1991), online book retailers (Latcovich and Smith, 2001), chemical manufacturing (Marin and Siotis, 2007), supermarkets (Ellickson, 2007), banking (Dick, 2007), newspapers and restaurants (Berry and Waldfogel, 2010), and across industries (Robinson and Chiang, 1997; Sutton, 1998). These analyses have focused upon examining the relationship between concentration, captured by the ratio of firm to industry sales, and investments in either capacity (Marin and Siotis, 2007), product quality (Ellickson, 2007; Berry and Waldfogel, 2010), or advertising (Robinson and Chiang, 1997; Latcovich and Smith, 2001).

Equations (7) and (8) lead directly to the empirically testable hypotheses for the lower bound to R&D concentration. Specifically, an industry characterized by endogenous fixed costs in R&D should exhibit a lower bound to R&D concentration that is non-decreasing in market size whereas R&D concentration in exogenous fixed cost industries is decreasing in market size. Sutton (1991) derives a formal test for the estimation of the lower bound to concentration in an industry, based upon Smith (1985, 1994), in which the concentration ratio is characterized by the (extreme value) Weibull distribution. As Sutton (1991, 1998) identifies, it is necessary to transform the R&D concentration ratio R_1 such that the predicted concentration measures will lie between 0 and 1. Specifically, the R_n concentration measure is transformed according to:[2]

$$\tilde{R}_n = \ln\left(\frac{R_n}{1 - R_n}\right). \tag{9}$$

We follow the functional form suggested by Sutton for the lower bound estimation such that for some sub-market m, the R_n concentration ratio is characterized by:

$$\frac{\tilde{R}_{nm}}{h_m^2} = \theta_0 + \theta_1 \frac{1}{h_m \ln(\Pi_m / F_0)} + \varepsilon_m, \tag{10}$$

where h_m is the degree of product heterogeneity, F_0 is the fixed setup cost, Π_m is total industry sales, and (θ_0, θ_1) are the parameters of the empirical model. The intercept parameter θ_0 reflects the theoretical lower bound as the market size becomes large whereas the slope parameter θ_1 reflects how the lower bound changes with changes in market size. The residuals ε between the observed values of R&D concentration and the lower bound are distributed according the Weibull distribution such that:

$$F(\varepsilon) = 1 - \exp\left[-\left(\frac{\varepsilon - \mu}{\delta}\right)^{\gamma}\right], \quad \gamma > 0, \delta > 0 \tag{11}$$

on the domain $\varepsilon \geq \mu$. The case of $\mu = 0$ corresponds to the two parameter Weibull distribution such that nonzero values of the shift parameter μ represent horizontal shifts of the distribution. The shape parameter γ corresponds to the degree of clustering of

[2] As the transformed R&D concentration is undefined for values of $R_{nm} = 1$, we monotonically shift the R&D concentration data by -0.01 prior to the transformation.

observations along the lower bound whereas the scale parameter δ captures the dispersion of the distribution.

To test for a lower bound to R&D concentration, it is equivalent to testing whether the residuals fit a two or three parameter Weibull distribution, that is testing whether $\mu = 0$. However, as Smith (1985) identifies, fitting equation (10) directly via maximum likelihood estimation is problematic for shape parameter values $\gamma \leq 2$.[3] Smith (1985, 1994) provides a two-step procedure for fitting the lower bound that is feasible over the entire range of shape parameter values.

Following the methodology of Giorgetti (2003), we first solve a linear programming problem using the simplex algorithm to obtain consistent estimators of $\{\theta_0, \theta_1\}$ in which the fitted residuals are non-negative. Therefore, $\{\hat{\theta}_0, \hat{\theta}_1\}$ solves:

$$\begin{aligned} &\min_{\{\theta_0,\theta_1\}} \sum_{m=1}^{N} \left[\frac{\tilde{R}_{nm}}{h_m^2} - \left(\theta_0 + \theta_1 \frac{1}{h_m \ln(\Pi_m/F_0)}\right)\right] \\ &s.t. \quad \frac{\tilde{R}_{nm}}{h_m^2} \geq \left(\theta_0 + \theta_1 \frac{1}{h_m \ln(\Pi_m/F_0)}\right), \forall m. \end{aligned} \qquad (12)$$

From the first step, we obtain parameter estimates for $\{\hat{\theta}_0, \hat{\theta}_1\}$ fitted residual values $\hat{\varepsilon}$ which can be used to estimate the parameters of the Weibull distribution via maximum likelihood. Specifically, as there are k parameters to be estimated in the first stage, there will be $N - k$ fitted residuals with positive values. By keeping only the fitted residuals with strictly greater than zero values, we maximize the log pseudo-likelihood function:

$$\max_{\{\gamma,\delta,\mu\}} \sum_{m=1}^{N-k} \ln\left[\left(\frac{\gamma}{\delta}\right)\left(\frac{\varepsilon-\mu}{\delta}\right)^{\gamma-1} \exp\left[-\left(\frac{\varepsilon-\mu}{\delta}\right)^{\gamma}\right]\right] \qquad (13)$$

[3] Specifically, for $1 < \gamma \leq 2$, a maximum for the likelihood function exists, but it does not have the same asymptotic properties and may not be unique. Moreover, for $0 \leq \gamma \leq 1$, no local maximum of the likelihood function exists.

with respect to $\{\gamma, \delta, \mu\}$ in order to test whether $\mu = 0$, which is equivalent to testing the two parameter versus three parameter Weibull distribution via a likelihood ratio test. If the three parameter Weibull distribution cannot be rejected, then this implies the presence of a horizontal shift in the distribution corresponding to an industry in which R&D is an exogenously determined sunk cost. In all cases, the likelihood ratio test fails to reject that the data fits the restricted, two parameter model such that $\mu = 0$. For each estimation, we report the likelihood ratio statistic which is distributed with a chi-squared distribution with one degree of freedom. Finally, we compute standard errors for the first-stage estimations via bootstrapping and standard errors for the second-stage estimations according to the asymptotic distributions defined in Smith (1994).

IV. Data and Descriptive Statistics

In order to estimate an endogenous fixed cost model à la Sutton (1991, 1998), it is necessary to have both firm-level sales data and total market size for each market that is representative of the entirety of the industry. Although such data are of limited availability for the agricultural biotechnology sector, estimation of the endogenous lower bound to R&D concentration in agricultural biotechnology according to the proposed model is feasible using publicly available data. The model specifically requires four types of data for each crop type: (i) firm-level data on R&D investment, (ii) industry-level data on sub-market size, (iii) industry-level data on product heterogeneity, and (iv) industry-level data on the minimum setup costs. Moreover, additional data on agricultural characteristics at the state level are required in order to separate the agricultural biotechnology sector into distinct sub-markets for each crop type.

Measuring R&D Concentration

For the empirical analysis of the agricultural biotechnology sector, we utilize two dimensions of variation in R&D investment and market size by estimating the lower bound across geographic sub-markets as well as over time. In doing so, we are able to capitalize upon changes in farmer and consumer attitudes towards GM crops over time as well as advances in technology and/or regulation which decrease the fixed costs associated with R&D. Moreover, geographic and inter-temporal variation in market size permits the theory to be tested across a variety of market sizes.

The ideal data for the analysis of an endogenous lower bound to R&D concentration would be R&D expenditures for each product line for every firm in an industry. Although data at this level of detail are unavailable for the agricultural biotechnology sector, there are publicly available data that capture proxies for R&D investment at the firm and product level in the form of patent and/or field trial applications for GM crops. However, data on crop patent applications are not available for the years after 2000 and therefore is less useful for an estimation of lower bounds to concentration for an industry in which there has been considerable consolidation post-2000. Field trial application (FTA) data are appropriate for the analysis as they capture an intermediate R&D process which is mandatory for firms that desire to bring a novel GM crop to market.

In accordance with the Federal Coordinated Framework for the Regulation of Biotechnology, the Animal and Plant Health Inspection Services (APHIS) regulates the release of any genetically engineered (GE) organism that potentially threatens the health of plant life. Specifically, prior to the release of any GE organism, the releasing agency, either firm or non-profit institution, must submit a permit application to the Biotechnology

Regulatory Services (BRS). (BRS, 2010) These Field Trial Applications (FTA) are made publicly available by the BRS in a database that includes information on all permits, notification, and petition applications for the importation, interstate movement, and release of GE organisms in the US for the years 1985-2010. The database includes the institution applying for the permit, the status of the application, the plant (or "article") type, the dates in which the application was received, granted, and applicable, the states in which the crops will be released, transferred to or originated from, and the crop phenotypes and genotypes. As of October 2010, there are 33,440 permits or notifications of release included in the database for all types of crops. After restricting the sample to firms, by eliminating non-profit institutions, and permits or notification involving the release of GE crops, there are 9,936 remaining observations in the database.

To our knowledge, there have been three previous studies that utilize FTA data to analyze an economic question. Most closely related to our analysis, Brennan, Pray, and Courtmanche (2000) find higher levels of R&D investment and increasing output among larger firms and ambiguous impact of merger and acquisition activity upon innovation. Stein and Rodríguez-Cerezo (2009) examine the commercialization potential of GM crops using FTA data and find evidence supporting the correlation between applications and commercialization especially for corn and soybean varieties. O'Connor (2009) examines the impact of differences in labeling regimes between the USA and the EU upon the incentives of firms to engage in R&D as captured by FTA data.

Figure 3 plots the annual number of field trial applications by private enterprises for each of the crop types between 1988 and 2010 and the yearly number of firms that submitted an application for each crop type. The data reveal that the number of distinct firms applying for at least one field trial in a given year peaked in the early 1990s for both cotton and soybean varieties and in 1995 for corn varieties before falling and roughly leveling off by 2000 for all varieties. This is consistent with an industry characterized by a new technology in which significant initial entry occurred, followed by a "shakeout" of firms when the commercialization possibilities of the technology are realized. The FTA data indicate that both GM corn and GM cotton applications increased throughout the 1990s, peaked in the early 2000s, and before falling and seemingly leveling off by 2010. The pattern for GM soybean applications reveals more gradual growth with a "peak" only appearing after 2005. The number of yearly applications, with the firm data, illustrates that average per firm R&D activity increased throughout the 1990s and leveled off in the 2000s, a story that is consistent with the slow introduction of subsequent generations of GM crops.

We also consider a measure of R&D concentration from field trial application data which has been adjusted for past mergers and acquisitions within each period. Specifically, if intellectual property assets are becoming increasingly concentrated among a smaller number of firms, then it is possible that the lower bound to R&D concentration is higher not due to endogenous fixed costs, but rather due to this consolidation activity. We utilize company histories and Lexis-Nexis news releases to identify merger and acquisition activity and the effective merger date in order to construct a measure of R&D concentration accounting for ownership changes. Although completed independently, our list of changes in corporate ownership corresponds to the activity reported in Fuglie, et al. (2011).

The Market for Agricultural Biotechnology

In estimating an EFC-type model for a single industry, an initial crucial step is the proper identification of the relevant product markets. The EFC model predicts an escalation of fixed-cost expenditures for existing firms as market size increases rather than entry by additional competitors. For the case of retail industries, such as those examined by Ellickson (2007) and Berry and Waldfogel (2010), markets are clearly delineated spatially. However, the identification of distinct markets in agricultural biotechnology is potentially more problematic as investments in R&D may be spread over multiple geographic retail markets. Moreover, as we only have data on firm concentration available at the state level, the difficulty associated with defining relevant markets is exacerbated.

In order to overcome issues associated with the correct market identification, we assume that R&D expenditures on GM crops released domestically can only be recouped on sales within the US. Although somewhat innocuous for the market for corn seed, this assumption may be overly restrictive for other crop types including soybeans and cotton. However, disparate regulatory processes across countries, as well as the significant size of the US market, reveals the importance of the domestic market to seed manufacturers. Recent surveys of global agricultural biotechnology indicate that many of the GM crop varieties adopted outside of the US have also been developed abroad. (James, 2010)

We consider a characterization of regional sub-markets for each crop variety derived from statistical cluster analysis of observable characteristics of agricultural production within each state and crop variety. Cluster analysis is a useful tool in defining regional sub-markets as it captures the "natural structure" of the data across multiple characteristics. The observable data used in the cluster analysis are from the period prior to the widespread adoption of GM varieties (1990-1995) and covers agricultural production in all lower, contiguous 48 states (except Nevada), although the extent of coverage varies by crop and state. The cluster analysis uses data, summarized in Table 1, that can be broadly classified into two types: (i) state-level data that are constant across crops; and (ii) data that vary by state and crop level. We utilize K-means clustering by minimizing the Euclidean distance of the observable characteristics for each crop variety and arrive at ten corn clusters, six soybean clusters, and six cotton clusters.[4] Results from the clustering analysis are reported in Table 2 along with the 2010 market shares for each sub-market.

Measuring Market Size, Product Heterogeneity, and Minimum Setup Costs

The National Agricultural Statistics Service (NASS), a division of the United States Department of Agriculture (USDA), conducts the annual June Agriculture Survey in order to obtain estimates of farm acreage for a variety of crops, including corn, cotton, and soybeans. The NASS reports data on total amount of acreage, both planted and harvested, in an annual *Acreage* report that is made publicly available. Moreover, the Economic Research Service (ERS) also computes yearly seed costs in dollars per acre based upon survey data collected by the USDA in the crop-specific Agricultural Resource Management Surveys (ARMS). After adjusting for inflation, these seed costs are multiplied by the total acres planted for each crop type to arrive at total market size which we use as a proxy for total industry sales in equation (10).

[4] For additional discussion of the cluster analysis of the geographically distinct sub-markets, please refer to Appendix A: (Sub-)Market Cluster Analysis for GM Crops.

Since 2000, the June Agricultural Survey has also sampled farmers regarding the adoption of GM seed varieties for corn, cotton, and soybeans across a subsample of states.[5] Using these survey data, ERS computes and reports estimates for the extent of GM adoption separated by crop type and GM characteristics for a sample of states. GM adoption rates, supplemented with the adoption data from Fernandez-Cornejo and McBride (2002) for the years 1996-1999, are used to obtain total GM acreage planted, as well as total market size after multiplying by the inflation adjusted dollar cost of seed, to arrive at a measure of GM market size for the lower bound estimation.

Additionally, the rates of adoption for 2000-2010, as well as the estimates of GM adoption for the years 1996-1999 from Fernandez-Cornejo and McBride (2002), are used to construct product heterogeneity indexes for each crop type that vary across time. By definition, the product heterogeneity index is meant to capture the percentage of industry sales of the largest product group. We treat seed varieties as homogenous within product groups, broadly defined as conventional, insect resistant (IR), herbicide tolerant (HT), and "stacked" varieties consisting of IR and HT traits, and equate the product heterogeneity index to the percentage of acres accounted for by the largest group. By constructing the heterogeneity index in this manner, we are likely to be introducing measurement error as the broad categories are likely to contain a variety of sub-categories corresponding to specific more specific traits. However, the potential measurement error that is introduced by this specification would actually bias our results towards the alternative hypothesis that GM seed markets are characterized by exogenous fixed costs.

---

[5] NASS estimates that the states reported in the GM adoption tables account for 81-86% of all corn acres planted, 87-90% of all soybean acres planted, and 81-93% of all upland cotton acres planted. For states without an adoption estimate, overall US adoption estimates are used to compute the size of the GM market.

The final component required for the estimation of the lower bound to R&D concentration is the minimum setup cost associated with product market entry. We use data reported in Frey (1996) and Traxler, et al. (2006) in order to obtain a proxy for the R&D setup cost for each crop type. The minimum setup cost is obtained by first summing the total number of public "scientist years" (SY), those reported by the State Agricultural Experiment Stations (SAES) and the Agricultural Research Service (ARS), and then dividing this sum by the total number of projects reported for both agencies to obtain an average SY for a single crop.[6] Minimum setup costs are obtained by multiplying the average SY by the private industry cost per SY ($148,000) and adjusting for inflation.[7]

V. Empirical Results and Discussion

Prior to estimating the lower bounds to R&D concentration, we examine whether the agricultural biotechnology sector appears to be characterized by an endogenous lower bound to R&D concentration. In Figure 4, the one- and four-firm R&D concentration ratios are plotted against the market size of each sub-market and for each crop type. Figure 4 illustrates that R&D concentration ratios are non-decreasing in market size for each crop type, regardless of the measure of R&D concentration, implying the possibility of a lower bound. However, these descriptive illustrations do not account for differing levels of product heterogeneity across time and therefore it is not possible to reconcile these illustrations directly with the lower bound to R&D concentration implied by the theory.

---

[6] A "scientist year" is defined as "work done by a person who has responsibility for designing, planning, administering (managing), and conducting (a) plant breeding research, (b) germplasm enhancement, and (c) cultivar development in one year (i.e., 2080 hours)."
[7] As a robustness check, we consider an alternate definition based upon public sector cost per SY ($296,750).

After adjusting for R&D concentration in each submarket according to equation (9), we estimate the relationship between the single firm R&D share and the ratio of total market size to private setup costs for each crop type. The subsequent section considers various robustness checks including alternate definitions of our explanatory variable, an alternate functional form for the first stage estimation, and an assumption of product homogeneity. In the final section, we account for merger and acquisition activity in the agricultural biotechnology sector in order to examine the impact of this consolidation activity upon concentration in R&D activities.

Estimating the Lower Bounds to R&D Concentration

The two-stage estimation results for each crop type are reported in Table 3 and Figures 5-7 illustrate the estimated lower bounds and 95% confidence intervals. The logit transformation of R&D concentration implies that a direct interpretation of the estimated coefficients is difficult. Specifically, the coefficient on the intercept term $\hat{\theta}_0$ can be interpreted as the theoretical lower bound as market size becomes large (i.e., $\frac{\Pi_m}{F_0} \rightarrow \infty$). The coefficient on market size $\hat{\theta}_1$, adjusted for product heterogeneity and fixed setup costs, indicates whether the theoretical lower bound is increasing (i.e., negative parameter), decreasing (i.e., positive parameter), or unchanging in market size. The null hypothesis of exogenous fixed costs implies that the theoretical bound to R&D concentration converges to 0 as market size becomes large (i.e., $\hat{\theta}_0 \approx -9.210$) and that the predicted bound is non-increasing in market size (i.e., $\hat{\theta}_1 \geq 0$).

The results from the first-stage estimations reveal that the predicted lower bound to R&D concentration is increasing in market size (i.e., $\hat{\theta}_1 < 0$) and a theoretical bound to R&D

concentration that is significantly different from zero across crop types. These results are consistent with an endogenous lower bound to R&D concentration as illustrated in Figure 2 in which concentration is very low in small-sized markets and increasing in market size. In order to derive the predicted theoretical lower bound for infinitely-sized markets, we multiply the theoretical lower bound parameter $\hat{\theta}_0$ by the square of the level of product heterogeneity in the largest sub-market for each crop type (i.e., corn: h = 0.479; cotton: h = 0.665; soybean: h = 0.922) and then perform an inverse logit transformation.

The theoretical lower bounds for each seed variety, reported in Table 3, illustrate that the largest firm in an infinitely-sized market would account for approximately 54.8% of R&D in corn seed, 47.3% of R&D in cotton seed, and 78.6% of R&D in soybean seed. Comparing the theoretical lower bounds to the fitted lower bounds for the currently largest-sized markets reveals that although the corn and cotton seed markets are already characterized by concentration in R&D, 33.0% and 35.9% respectively, there is still significant room for additional concentration. Although R&D is heavily concentrated in many of the sub-markets for soybean seeds, for the largest-sized current market the predicted share of R&D is only 20.8%.

Figures 5-7 graph the theoretical and fitted lower bounds to R&D concentration for each seed variety as well as the current levels of R&D concentration and 95% lower bound confidence intervals. The clustering of observations between the theoretical and fitted lower bounds in GM corn seed (Figure 5) illustrates a convergence to the theoretical lower bound as markets become large consistent with endogenous fixed costs. The relationship for GM cotton seed (Figure 6) and GM soybean seed (Figure 7) between market size and concentration is less straightforward. Part of the difficulty in identifying endogenous fixed costs in these markets arises from the relatively small number of GM seed varieties (i.e., 11 GM cotton varieties and 5 GM soybean varieties) for these crop types. (Howell, et al., 2009)

It is also important to address the similarities and differences in the second-stage estimations for GM corn, cotton, and soybean seeds. From the second stage results, the likelihood ratio tests with one degree of freedom fail to reject the null hypothesis that the first stage residuals fit a two-parameter Weibull distribution for all crop types and market definitions. Hence, we fail to reject the hypothesis that the markets for GM corn, cotton, and soybean are characterized by endogenous fixed costs.

Recall that the parameter γ corresponds to the shape of the Weibull distribution such that a lower value of γ corresponds to a higher degree of clustering around the lower bound and the scale parameter δ describes the dispersion of the data. The results on the shape parameter γ imply a high degree of clustering on the lower bound for all crop types, with corn the most clustering. Moreover, γ is less than two in all estimations implying that the two-step procedure of Smith (1985, 1994), is appropriate. Finally, the estimations of the scale parameter δ indicate a wider dispersion of R&D concentration in corn sub-markets relative to cotton and soybean sub-markets.

Estimating the Lower Bounds to R&D Concentration: Robustness Checks

We consider four separate robustness checks for our estimations of the lower bound to R&D concentration for GM seeds. First, we alter the definition of the market size for GM crops by measuring the market size for GM seed varieties only rather than the market size for all varieties. The second robustness check examines the importance of the heterogeneity index by assuming homogenous products. The theory implies that this extreme assumption should yield lower bound estimations consistent with exogenous fixed costs and allows us to examine the importance of product differentiation for each of these GM crop types. The third robustness check considers an alternate specification of minimum setup costs using public costs rather than private costs. Finally, the last specification considers an alternate functional form by including a quadratic market size term. The robustness checks, reported in Table 4, generally confirm our results of endogenous fixed costs for all seed varieties with a few interesting exceptions.

As a first robustness check, we consider an alternate definition of market size by defining the relevant market according to the actual number of acres planted with GM varieties for each crop type. This definition deviates from our previous measure of market size in two dimensions: first, we drop the observations from the first period (1991-1995) since there were no commercial GM varieties available in this time period; and second, the total number of acres planted for each crop type is measured according to the percent of acres planted with GM varieties. For GM cotton seeds and GM soybean seeds, this robustness check supports our finding of endogenous fixed costs with the same sign and similar magnitudes for our estimates of the market size parameter $\hat{\theta}_1$ and greater theoretical lower bound estimates.

The results for the first robustness check, estimating the lower bound to concentration using the alternate definition of market size, are reported in Table 4 in the "GM Market Size" columns for each seed variety. The results for cotton and soybean seeds under the alternate definition of market size support our findings of endogenous fixed costs to R&D with a theoretical lower bound to concentrations under the alternate definition which are greater (i.e., 0.600 vs. 0.473 for cotton seeds and 0.953 vs. 0.786 for soybean seeds) and a fitted lower bound to market size that increases more rapidly under the alternate definition of market size.

For the alternate definition of market size for GM corn seeds however, the results appear to provide evidence against endogenous fixed costs in favor of exogenous fixed costs. Although the second-stage estimates continue to support endogenous fixed costs by rejecting the three-parameter Weibull specification for the distribution of residuals, the first-stage estimates imply that the fitted lower bound is decreasing in the size of the GM market with a theoretical lower bound of only 11.9% share of R&D for the market-leading firm. However, an examination of the underlying data reveals that these results are being driven by an outlying sub-market (Northeast states) which has a very small market size and relatively concentrated R&D. The estimation eliminating the outlying sub-market results in parameter estimates which are comparable to our other estimates under the total market size definition in both size and magnitude, albeit with a lower bound that increases more gradually with market size and a lesser theoretical lower bound.

The second robustness check estimates the lower bound to concentration under the assumption that seed varieties within each crop type are homogenous and are reported in Table 4 under the "Homogeneity" column. Theory predicts that under perfect product homogeneity, industries will be characterized by exogenous fixed costs as firms are unable to differentiate their products and gain increased market share. We find that under the assumption of perfect homogeneity, the lower bound to concentration across all three crop varieties is characterized by exogenous fixed costs (i.e., fitted lower bounds that are non-increasing in market size). These results confirm the importance of product heterogeneity in the sector and imply that if firms fail to innovate and differentiate their varieties, we should expect decreasing levels of concentration in R&D.

The next robustness check considers the possibility of measurement error in the minimum setup cost data and the results are reported in Table 4 under the "Setup Cost" column. If we are undervaluing the true measure of minimum setup costs in our estimations, then our estimations could be biased towards finding endogenous fixed costs. We consider an alternate definition, public cost of science years rather than private cost of science years, which increases the minimum setup cost for each crop. Under this specification of setup costs, we continue to find evidence of endogenous fixed costs across crop types and our parameter estimates move in the predicted direction. Specifically, the fitted lower bound increases more gradually with changes in market size, the coefficients on the adjusted market size parameter are smaller in absolute value, and the theoretical lower bounds are (insignificantly) smaller.

Finally, we consider an alternate specification of the lower bound to R&D concentration which permits the lower bound to change non-linearly with market size by including both adjusted market size and the square of adjusted market size. The estimation results for GM corn and soybean seeds, reported in Table 4 under the "Quadratic Market Size" columns, continue to provide support for the existence of endogenous fixed costs. Specifically, the fitted lower bound to concentration is increasing with the first-order effect of market size ("Adjusted Market Size") at a decreasing rate ("Adjusted Market Size Squared"). However, the parameter estimates for GM cotton seed under this alternate functional specification now imply that this crop type is characterized by exogenous fixed costs (i.e., a fitted lower bound to concentration that is decreasing in market size).

However, an examination of the underlying data reveals a single outlier (Western states in 2006-2010 characterized by a small market and low levels of concentration) which is driving the estimation results. The results from unreported estimations in which we eliminate this single outlying observation confirm the presence of endogenous fixed costs.

Estimating the Lower Bounds to R&D Concentration under Firm Consolidation

The final set of estimations, reported in Table 5, consider the impact of mergers and acquisitions in the agricultural biotechnology sector upon R&D concentration. For the corn seed and cotton seed varieties, we reject the null hypothesis of exogenous fixed costs with the fitted lower bound increasing in market size and a theoretical lower bound that is significantly different from zero. However, the results for the lower bound to R&D concentration in the market for soybean seeds changes dramatically after accounting for merger and acquisition activity and we cannot reject the hypothesis of exogenous fixed costs. Figures 8-10 plot the fitted lower bound and 95% confidence intervals for both the baseline estimations and the estimations adjusted R&D concentration for merger and acquisition activity. Additionally, in Table 5 we test the equivalence of the first-stage parameter estimates in order to determine whether merger and acquisition activity has significantly increased or decreased the bounds to R&D concentration.

For GM corn seed, Figure 8 illustrates that adjusting R&D concentration for merger and acquisition activity does not appear to change the fitted lower bound to R&D concentration. However, the results in Table 5 indicate that the theoretical lower bound, represented by the intercept parameter $\hat{\theta}_0$, is significantly greater after accounting for mergers and acquisitions whereas the convergence to this bound, represented by the market size parameter $\hat{\theta}_1$, is not significantly different. Although the theoretical lower bound under mergers and acquisitions is significantly different, the economic significance is negligible as the R&D share of the leading firm increases from 54.8% to only 56.3%.

Contrary to intuition that merger and acquisition activity raises the lower bound to R&D concentration, the estimates for GM soybean seed imply that the consolidation activity has resulted in this market to be characterized by exogenous fixed costs. The results in Table 6 of the test of the equivalence of parameter estimates reveal that both the theoretical lower bound parameter $\hat{\theta}_0$ and the market size parameter $\hat{\theta}_1$ are significantly different, after accounting for mergers and acquisitions. Figure 10 illustrates the difference in the fitted bounds and, comparing the plotted concentration measures to those in Figure 7, reveals that the difference in these results can be largely attributed to increased R&D concentration in medium-sized sub-markets.

Finally, we fail to find any significant difference in R&D concentration, measured by either the fitted lower bound or the theoretical lower bound, after accounting for firm mergers and acquisitions in the cotton seed market. Overall, there is little evidence to support the claim that merger and acquisition activity in the agricultural biotechnology sector has significantly increased the concentration of intellectual property assets in excess of the endogenous fixed cost nature of innovation.

VI. Conclusions

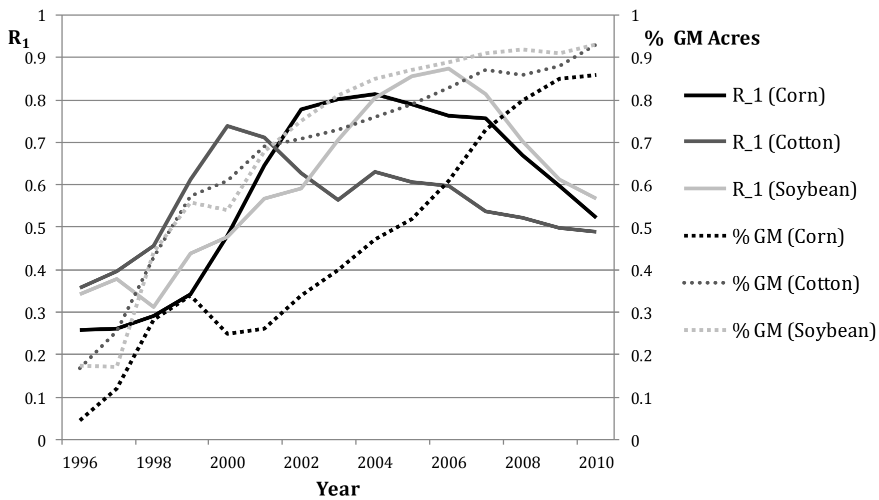
We examine whether the agricultural biotechnology sector is characterized by endogenous fixed costs (EFC) associated with R&D investment. In a mixed model of vertical and horizontal product differentiation, we derive the theoretical lower bound to R&D

concentration from Sutton's (1998) EFC model of market concentration and innovation. Using data on field trial applications of genetically modified (GM) crops, we estimate the lower bound to R&D concentration in the agricultural biotechnology sector. We identify the lower bound to concentration using exogenous variation in market size across time, as adoption rates of GM crops increase, and across agricultural regions.

The results of the empirical estimations imply that the markets for GM corn, cotton, and soybean seeds are characterized by endogenous fixed costs associated with R&D investments. In particular, the leading corn seed producer has an estimated R&D share, within an infinitely-sized market, of 54.8%. This theoretical lower bound is lower for the leading cotton seed producer (47.3%) and is higher for the largest soybean seed producer (78.6%). Moreover, adjusting for firm consolidation via mergers and acquisitions does not significantly raise the lower bound estimations for either the cotton or soybean markets. Although the lower bound to R&D concentration in GM corn seed is statistically greater after accounting for merger and acquisition activity, the difference (56.3% with M&A and 54.7% without M&A) is not economically significant. These results imply that concerns over the concentration of intellectual property, as measured by field trial applications, resulting from mergers and acquisitions in agricultural biotechnology appear unfounded.

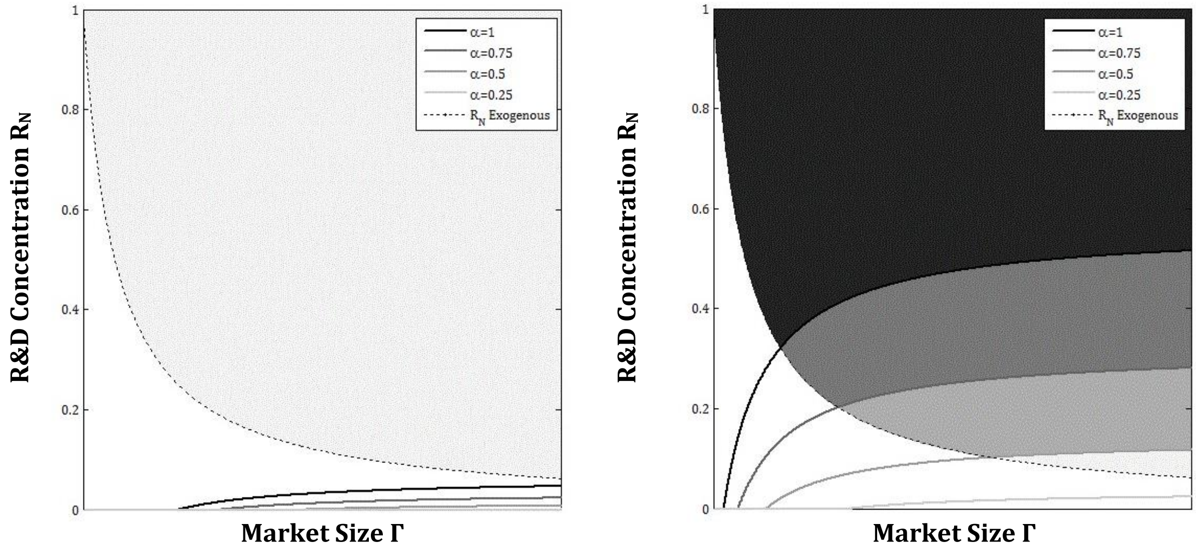
Given the increased concerns over concentration in agricultural inputs, and in particular in agricultural biotechnology, regulators and policymakers alike will find these results of particular interest. Whereas increased levels of concentration are often associated with anticompetitive behavior in an industry, the presence of endogenous fixed costs and the nature of technology competition in agricultural biotechnology imply a certain level of concentration is to be expected. Specifically, R&D activity is concentrated within three to four firms across corn, cotton, and soybeans and the ratios of concentration have been changing little over the past 20 years. Moreover, the empirical model leaves open the possibility that the introduction of second and third generation GM varieties, the opening of foreign markets to GM crops, future exogenous shocks to technology, or reductions in regulatory cost could lead to additional entry, exit, or consolidation in the industry.

Figure 1: Single-Firm R&D Concentration Ratios (R_1) and GM Adoption



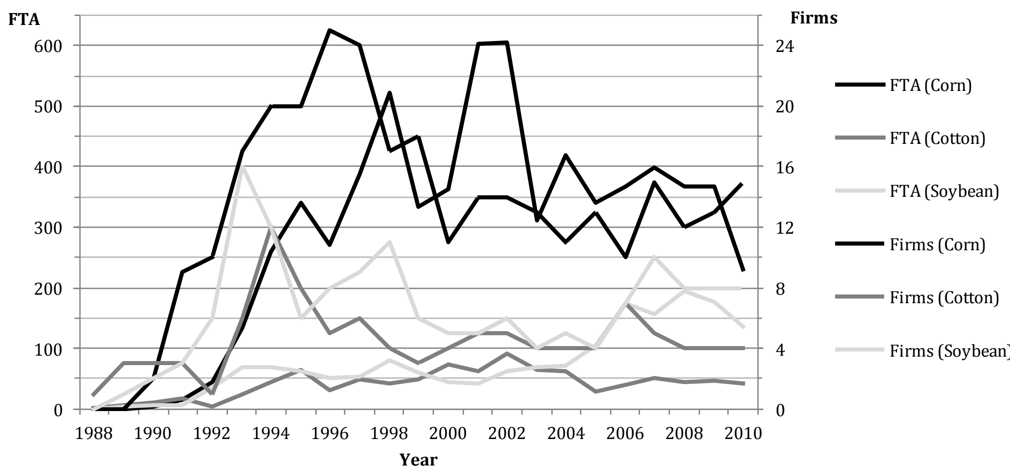
Source: Authors' calculations from APHIS data and Fernandez-Cornejo (2013).

Figure 2: Equilibrium R&D Concentration Levels and Market Size



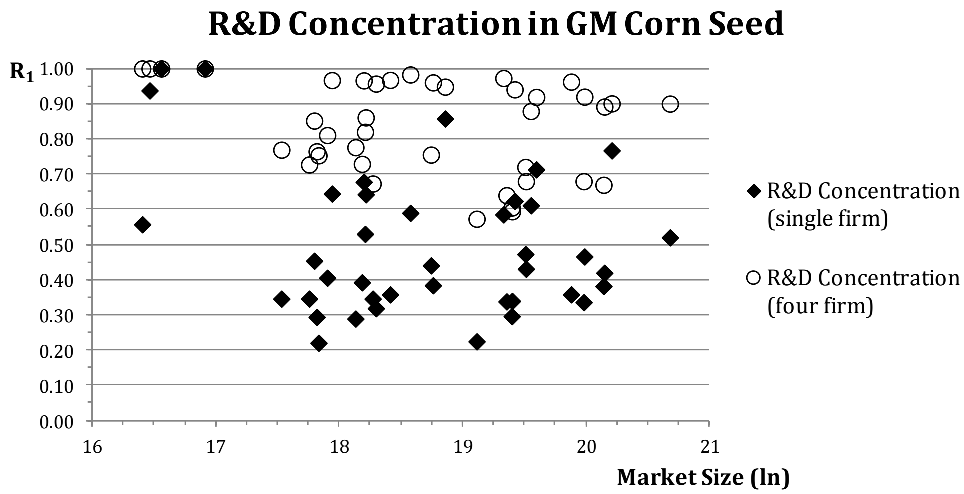
(a) Low *h* (*h* = 0.75) (b) High *h* (*h* = 0.25)

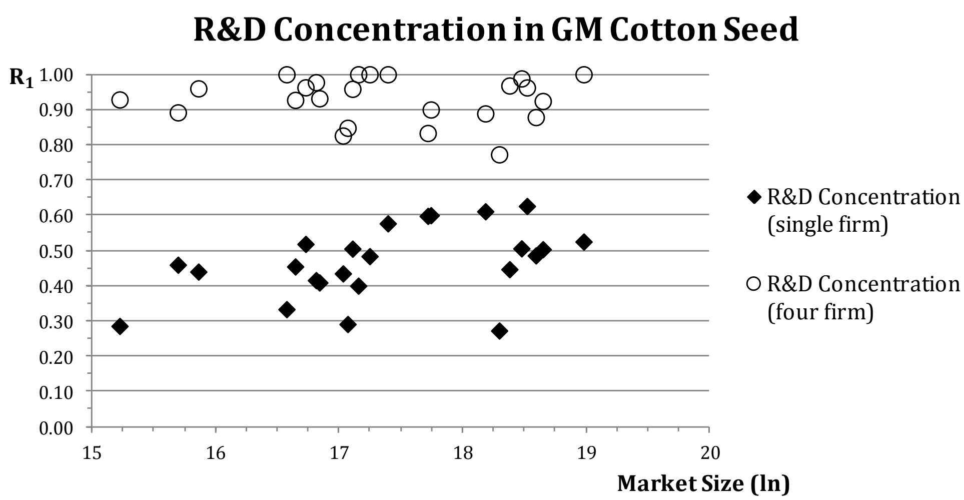
Figure 3: Field Trial Applications (FTA) and Number of Firms by Crop Type

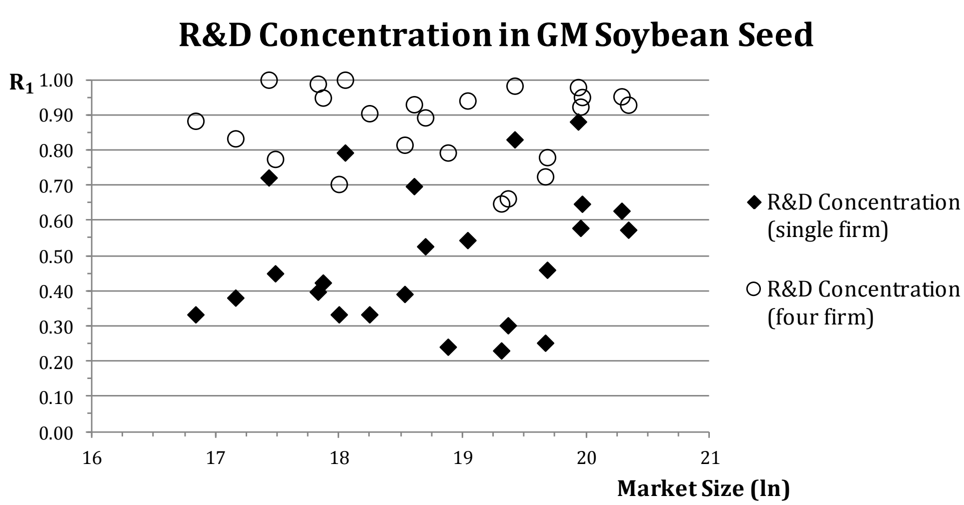


Source: Authors' calculations from APHIS data.

Figure 4: R&D Concentration and Market Size in Agricultural Biotechnology

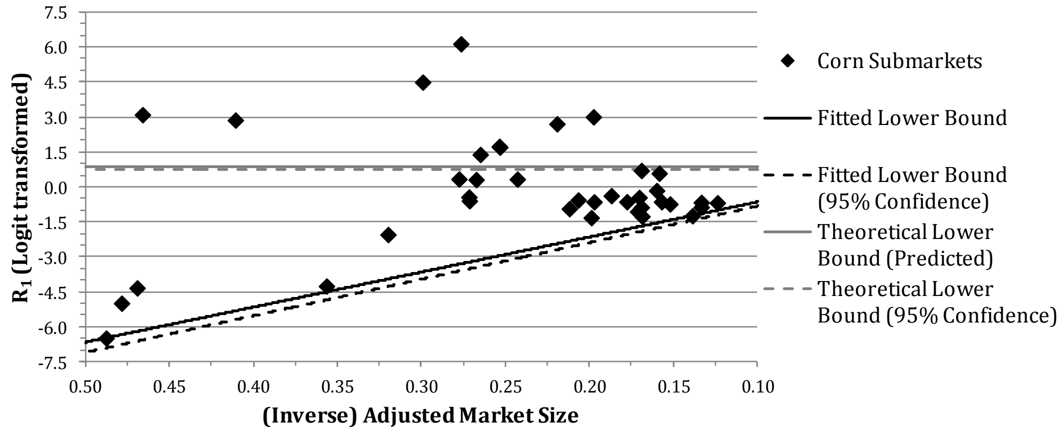






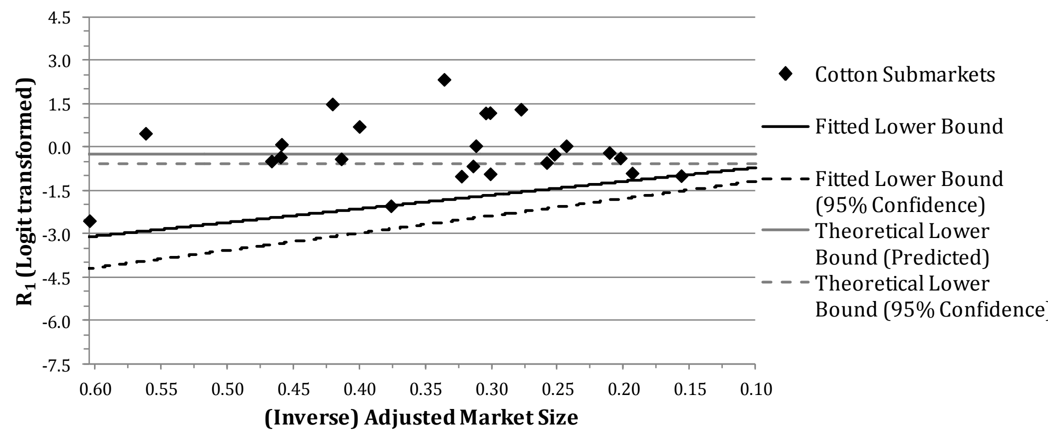
Source: Authors' calculations.

Figure 5: Lower Bound to R&D Concentration (GM Corn Seed)



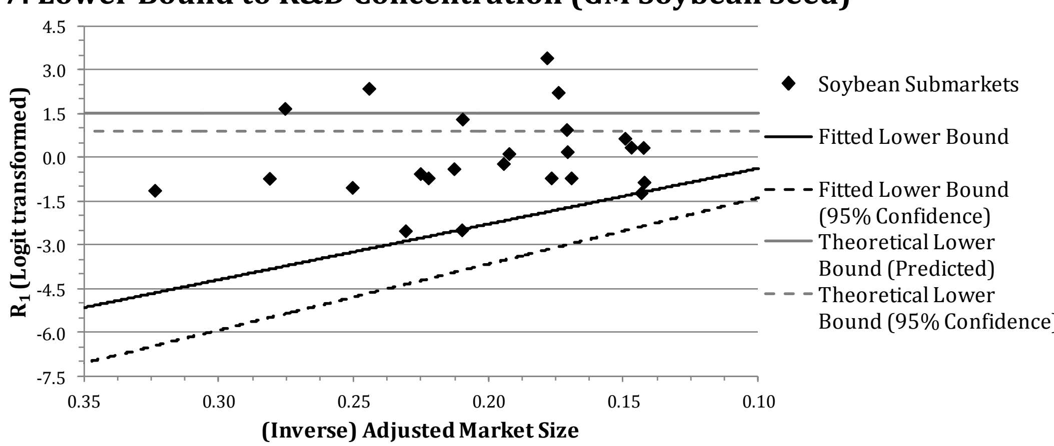
Source: Authors' calculations.
Note: Submarkets for Northeast states (2001-2005, 2006-2010) have high levels of R&D concentration and small market sizes and have been omitted from Figure 5 without loss of generality.

Figure 6: Lower Bound to R&D Concentration (GM Cotton Seed)



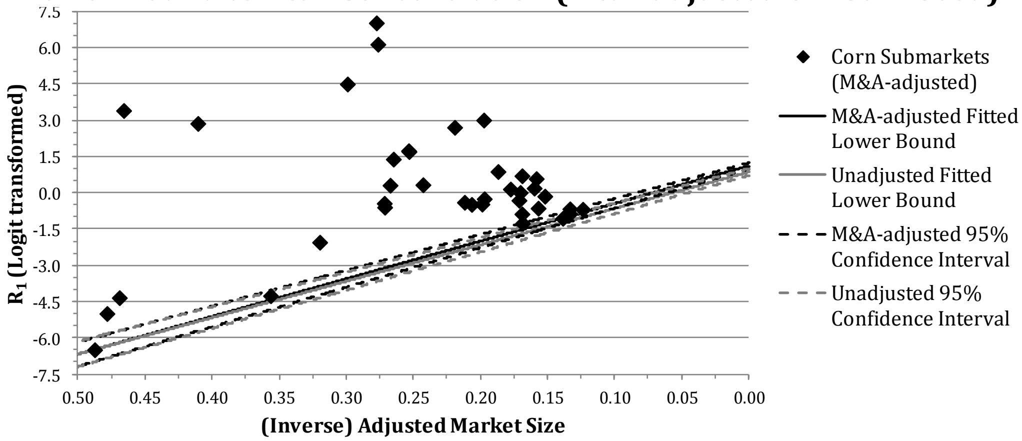
Source: Authors' calculations.

Figure 7: Lower Bound to R&D Concentration (GM Soybean Seed)



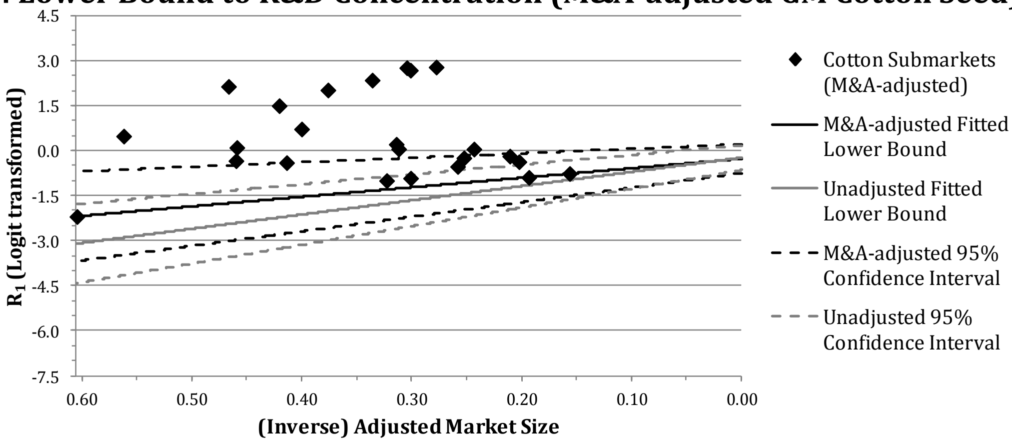
Source: Authors' calculations.

Figure 8: Lower Bound to R&D Concentration (M&A-adjusted GM Corn Seed)



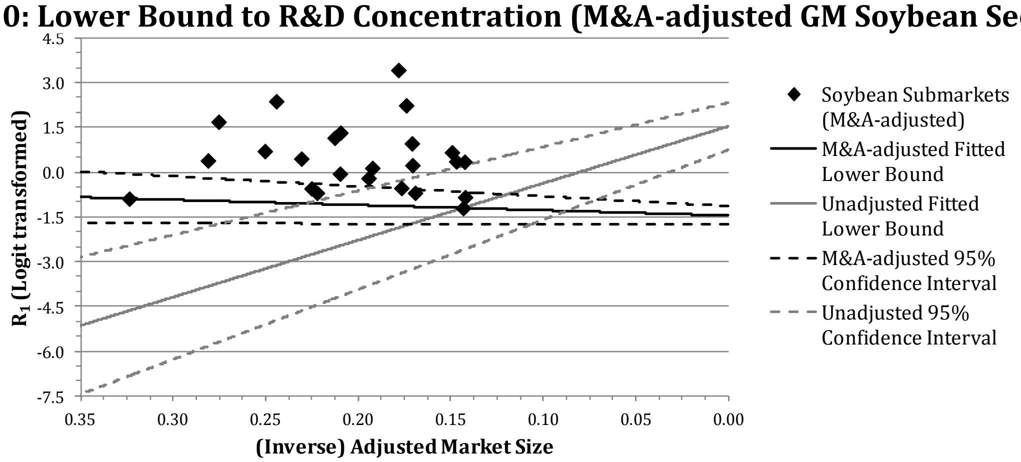
Source: Authors' calculations.
Note: Submarkets for Northeast states (2001-2005, 2006-2010) have high levels of R&D concentration and small market sizes and have been omitted from Figure 5 without loss of generality.

Figure 9: Lower Bound to R&D Concentration (M&A-adjusted GM Cotton Seed)



Source: Authors' calculations.

Figure 10: Lower Bound to R&D Concentration (M&A-adjusted GM Soybean Seed)



Source: Authors' calculations.

Table 1: Observable Market Characteristics for Sub-Market Analysis			
State Level			
Data	**Description**	**Years**	**Source**
Latitude	State geographic centroid	-	Rosenberg
Longitude	State geographic centroid	-	Rosenberg
Size	Total area (000s acres)	-	US Census State & County QuickFacts
Temperature	Monthly averages (°F)	1971-2000	NOAA
Rainfall	Monthly averages (inches)	1971-2000	NOAA
Drought Likelihood	Monthly averages (PDSI)	1971-2001	NOAA
R&D	Total public funds for agricultural R&D (1990 $000s)	1990-1995	CRIS
Cropland	Total cropland area (000s acres)	1987;1992	Census of Agriculture
State and Crop Level			
Data	**Description**	**Years**	**Source**
Acres Planted*	Total area planted (000s acres)	1987;1992	Census of Agriculture
Share of Cropland*	Percentage of cropland planted (%)	1987;1992	Census of Agriculture
Farms*	Total farms (farms)	1987;1992	Census of Agriculture
Average Farm Size*	Average farm size (000s acres)	1987;1992	Census of Agriculture
Farms with Sales*	Total farms selling (farms)	1987;1992	Census of Agriculture
Sales*	Total sales (1990 $000s)	1987;1992	Census of Agriculture
Average Farm Sales*	Average farm sales (1990 $000s)	1987;1992	Census of Agriculture
Fertilizer Usage (3 types)**	Percentage of planted acres treated (%)	1990-1995	Agricultural Chemical Usage
Herbicide Usage (All types)**	Percentage of planted acres treated (%)	1990-1995	Agricultural Chemical Usage
Insecticide Usage (All types)***	Percentage of planted acres treated (%)	1990-1995	Agricultural Chemical Usage

*: Corn - No NV; Cotton - Only AL, AZ, AR, CA, FL, GA, KS, LA, MS, MO, NM, NC, OK, SC, TN, TX, VA; Soybean - No AZ, CA, CT, ID, ME, MA, MT, NV, NH, NM, NY, OR, RI, UT, WA, WY

**: Corn - No NV; Cotton - Only AZ, AR, CA, LA, MS, TX; Soybean - No AZ, CA, CO, CT, ID, ME, MA, MT, NV, NH, NM, NY, OR, RI, UT, VT, WA, WV, WY

***: Corn - No NV; Cotton - Only AZ, AR, CA, LA, MS, TX; Soybean - Only AR, GA, IL, IN, KY, LA, MS, MO, NE, NC, OH, SD

Table 2: GM Seed Sub-Markets by Crop Type		
Sub-Market	**2010 Market Shares (%)**	**States**
<u>Corn Seed Markets</u>		
"Core" States	29.47	IL, IA
Western "Fringe" States	17.64	NE, ND, SD
Eastern "Fringe" States	15.22	KY, IN, MI, OH
Northern "Fringe" States	12.97	MN, WI
Southern "Fringe" States	9.10	KS, MO
Southern Plains / Mississippi Delta States	4.87	AR, LA, MS, OK, TX
Mid-Atlantic States	4.16	DE, MD, NJ, NY, PA, VA, WV
Southeastern States	2.98	AL, FL, GA, NC, SC, TN
Western States	3.38	AZ, CA, CO, ID, MT, NV, NM, OR, UT, WA, WY
Northeastern States	0.20	CT, ME, MA, NH, RI, VT
<u>Cotton Seed Markets</u>		
Texas	52.42	TX
Southeastern States	20.95	AL, FL, GA, SC, TN
Mississippi Delta States	10.82	AR, LA, MS
Atlantic States	5.87	NC, VA
Southern Plains States	5.04	KS, MO, OK
Southwestern States	3.19	AZ, CA, NM
<u>Soybean Seed Markets</u>		
Western "Core" States	31.53	IA, MN, MO, WI
Eastern "Core" States	28.91	IL, IN, KY, MI, OH
Northern Plains States	22.38	KS, NE, ND, SD
Southeastern States	9.40	AL, FL, GA, NC, SC, TN
Southern Plains / Mississippi Delta States	9.05	AR, LA, MS, OK, TX
Mid-Atlantic States	2.72	DE, MD, NJ, NY, PA, VA, WV

Source: Authors' estimates from NASS (2010) Acreage Report

Table 3: Lower Bound Estimations for GM Crops			
Adjusted R&D Concentration Ratio (R_1)	**Corn Seed**	**Cotton Seed**	**Soybean Seed**
First-Stage			
Adjusted Market Size (θ_1)	-15.045 **	-4.735 **	-19.075 **
	(0.380)	(0.719)	(2.080)
Intercept (θ_0)^	0.848 **	-0.241 **	1.531 **
	(0.069)	(0.200)	(0.381)
Second-Stage			
Shape Parameter (γ)	0.713 **	1.815 **	1.802 **
	(0.013)	(0.034)	(0.064)
Scale Parameter (δ)	4.076 **	2.142 **	2.872 **
	(0.160)	(0.057)	(0.076)
Theoretical Lower Bound (R_1^{∞})^^	**0.548**	**0.473**	**0.786**
Lower Bound (95% confidence)	0.544	0.436	0.681
Feasible Range ($h \in [0,1]$)	0.500-0.700	0.440-0.500	0.500-0.822
Fitted Lower Bound (Largest Submarket)	0.330	0.359	0.208
Likelihood Ratio (χ^2=1)	-0.137	0.001	0.017
First-stage Observations	40	24	24
Second-stage Observations	38	22	22

Source: Authors' estimates.

Standard errors in parentheses.

^: Null hypothesis (H_0): $\theta_0 \approx -9.210$.

Null hypothesis (H_0): As the market size becomes large, does the lower bound to R&D concentration converge to (approximately) 0 assuming homogeneity (h = 1)?

^^: Bounds calculated using product heterogeneity for largest submarket for each seed variety (corn: h = 0.479; cotton: h = 0.665; soybean: h = 0.922) and infinitely-sized markets.

**,*: Significance at the 99% and 95% levels, respectively.

Table 4: Robustness Checks on Lower Bound Estimations

	Corn Seed				Cotton Seed				Soybean Seed			
Adjusted R&D Concentration Ratio (R_1)	GM Market Size	Homogeneity (h = 1)	Setup Costs (↑ F_0)	Quadratic Market Size	GM Market Size	Homogeneity (h = 1)	Setup Costs (↑ F_0)	Quadratic Market Size	GM Market Size	Homogeneity (h = 1)	Setup Costs (↑ F_0)	Quadratic Market Size
First-Stage												
Adjusted Market Size (θ_1)	3.909 **	-0.760	-13.202 **	-34.686 **	-5.604 *	0.446	-3.875 **	12.241 **	-20.330 **	6.825 **	-17.375 **	-34.876 **
	(0.995)	(1.374)	(0.391)	(1.850)	(2.306)	(1.071)	(0.594)	(1.427)	(1.075)	(1.018)	(1.885)	(6.588)
Adjusted Market Size Squared (θ_2)	-	-	-	110.377 **	-	-	-	-64.333 **	-	-	-	113.046 **
				(9.853)				(6.389)				(22.367)
Intercept (θ_0)^	-5.737 **	-1.134 **	0.786 **	1.449 **	0.835 **	-1.051 **	-0.302 **	-1.322 **	2.999 **	-2.211 **	1.560 **	1.472 **
	(0.455)	(0.212)	(0.077)	(0.152)	(1.036)	(0.227)	(0.196)	(0.171)	(0.167)	(0.169)	(0.379)	(1.195)
Second-Stage												
Shape Parameter (γ)	0.946 **	1.062 **	0.708 **	0.726 **	1.691 **	2.620 **	1.843 **	1.573 **	1.221 **	1.282 **	1.823 **	1.389 **
	(0.022)	(0.019)	(0.012)	(0.013)	(0.043)	(0.022)	(0.034)	(0.031)	(0.060)	(0.048)	(0.065)	(0.053)
Scale Parameter (δ)	7.260 **	1.765 **	4.059 **	4.203 **	1.913 **	0.996 **	2.121 **	1.774 **	1.584 **	1.295 **	2.986 **	2.164 **
	(0.292)	(0.047)	(0.160)	(0.166)	(0.074)	(0.018)	(0.055)	(0.056)	(0.086)	(0.048)	(0.078)	(0.078)
Heterogeneity Index (h)	0.591	1.000	0.479	0.479	0.697	1.000	0.665	0.665	1.000	1.000	0.922	0.922
Theoretical Lower Bound (R_1^∞)^^	**0.119**	**0.244**	**0.545**	**0.582**	**0.600**	**0.259**	**0.467**	**0.358**	**0.953**	**0.099**	**0.790**	**0.778**
Lower Bound (95% confidence)	0.093	0.184	0.540	0.571	0.389	0.192	0.431	0.329	0.938	0.076	0.686	0.386
Feasible Range ($h \in [0,1]$)	0.003-0.500	-	0.500-0.687	0.500-0.810	0.500-0.697	-	0.425-0.500	0.210-0.500	0.500-9.530	-	0.500-0.826	0.500-0.813
Likelihood Ratio (χ^2=1)	-0.117	-0.413	-0.052	-0.176	-0.034	-0.226	0.000	-0.035	-0.072	0.018	0.016	0.015
First-stage Observations	30	40	40	40	18	24	24	24	18	24	24	24
Second-stage Observations	28	38	38	37	16	22	22	21	16	22	22	21

Source: Authors' estimates.

Standard errors in parentheses.

^: Null hypothesis (H_0): $\theta_0 \approx -9.210$.

Null hypothesis (H_0): As the market size becomes large, does the lower bound to R&D concentration converge to (approximately) 0 assuming homogeneity (h = 1)?

^^: Bounds calculated using product heterogeneity for largest submarket for each seed type and infinitely-sized market.

**,*: Significance at the 99% and 95% levels, respectively.

Table 5: Lower Bound Estimations for GM Crops (Mergers and Acquisitions Adjusted)			
Adjusted R&D Concentration Ratio (R_1)	**Corn Seed**	**Cotton Seed**	**Soybean Seed**
First-Stage			
Adjusted Market Size (θ_1)	-15.573 **	-3.173 **	1.752 *
	(0.377)	(0.806)	(0.763)
Intercept (θ_0)^	1.105 **	-0.276 **	-1.456 **
	(0.059)	(0.232)	(0.150)
Second-Stage			
Shape Parameter (γ)	0.884 **	1.232 **	1.664 **
	(0.013)	(0.032)	(0.058)
Scale Parameter (δ)	5.587 **	2.050 **	1.902 **
	(0.168)	(0.079)	(0.055)
Theoretical Lower Bound (R_1^{∞})^^	**0.563**	**0.470**	**0.225**
Lower Bound (95% confidence)	0.560	0.427	0.189
Feasible Range ($h \in [0,1]$)	0.500-0.751	0.431-0.500	.189-0.500
Fitted Lower Bound (Largest Submarket)	0.337	0.405	0.262
Likelihood Ratio (χ^2=1)	-0.053	-0.008	0.013
First-stage Observations	40	24	24
Second-stage Observations	38	22	22

Source: Authors' estimates.

Standard errors in parentheses.

^: Null hypothesis (H_0): $\theta_0 \approx -9.210$.

Null hypothesis (H_0): As the market size becomes large, does the lower bound to R&D concentration converge to (approximately) 0 assuming homogeneity (h = 1)?

^^: Bounds calculated using product heterogeneity for largest submarket for each seed variety (corn: h = 0.479; cotton: h = 0.665; soybean: h = 0.922) and infinitely-sized markets.

**,*: Significance at the 99% and 95% levels, respectively.

Table 6: Impact of Mergers and Acquisitions upon R&D Concentration						
	Corn Seed		Cotton Seed		Soybean Seed	
Parameter Equivalence						
$t$test: $\hat{\theta}_{0_{\mathrm{U}}} = \hat{\theta}_{0_{\mathrm{M\&A}}}$	2.554*		-0.030		7.429**	
$t$test: $\hat{\theta}_{1_{\mathrm{U}}} = \hat{\theta}_{1_{\mathrm{M\&A}}}$	-0.888		0.386		-9.318**	
$df(\hat{\theta}_0)$	64		46		30	
$df(\hat{\theta}_1)$	70		46		29	
Variance Equivalence						
Ftest: $s_{\mathrm{U}}^2(\hat{\theta}_0) = s_{\mathrm{M\&A}}^2(\hat{\theta}_0)$	2.822**		1.896		6.110**	
Ftest: $s_{\mathrm{U}}^2(\hat{\theta}_1) = s_{\mathrm{M\&A}}^2(\hat{\theta}_1)$	2.179**		1.688		7.113**	
	Unadj.	**M&A**	**Unadj.**	**M&A**	**Unadj.**	**M&A**
Fitted Values						
$\hat{\theta}_0$	0.848	1.105	-0.241	-0.276	1.531	-1.456
$\hat{\theta}_1$	-15.045	-15.573	-4.735	-3.173	-19.075	1.752
Standard Deviation						
$s(\hat{\theta}_0)$	0.541	0.322	0.949	1.306	1.787	0.723
$s(\hat{\theta}_1)$	3.075	2.084	3.418	4.440	10.036	3.763
Sample Variance						
$s^2(\hat{\theta}_0)$	0.293	0.104	0.900	1.706	3.195	0.523
$s^2(\hat{\theta}_1)$	9.459	4.342	11.682	19.715	100.730	14.161
Observations (N)	40	40	24	24	24	24

Source: Authors' estimates.

**,*: Significance at the 99% and 95% levels, respectively.

References

Berry, S. and Waldfogel, J. "Product Quality and Market Size." *The Journal of Industrial Economics*, Vol. 58 (2010), pp. 1-31.

Biotechnology Regulatory Services (BRS). "USDA-APHIS Biotechnology Regulatory Services User Guide: Notification." BRS, Animal and Plant Health Inspection Service (APHIS), United States Department of Agriculture (USDA), 2010.

Brennan, M.F., Pray, C.E. and Courtmanche, A. "Impact of Industry Concentration on Innovation in the U.S. Plant Biotech Industry." Proceedings: *Transitions in Agbiotech: Economics of Strategy and Policy*, Washington, DC, June 24-25, 1999.

Chen, Q., Naseem, A. and Pray, C. "The Determinants of Research Productivity in the Agricultural Input Industry." Prepared for *The American Agricultural Economics Association Annual Meeting*, Denver, CO, Aug. 1-4, 2004.

Department of Justice. "Justice Department and USDA to Hold Public Workshops to Explore Competition Issues in the Agriculture Industry." Press Release #09-771, U.S. Department of Justice, 2009.

Dick, A. "Market Size, Service Quality, and Competition in Banking." *Journal of Money, Credit, and Banking*, Vol. 39 (2007), pp. 49-81.

Dillon, M. and Hubbard, K. "Rebuttal to Monsanto's Comments to the Departments of Justice and Agriculture." Submitted in response to the Request for Comments by the United States Department of Agriculture and United States Department of Justice, Antitrust Division, in connection with their hearings on "Agricultural and Antitrust Enforcement Issues in Our 21st Century Economy". Organic Seed Alliance, 2010.

Ellickson, P. "Does Sutton Apply to Supermarkets?" *RAND Journal of Economics*, Vol. 38 (2007), pp. 43-59.

Fernandez-Cornejo, J. "Recent Trends in GE Adoption." Economic Research Service (ERS), US Department of Agriculture (USDA), 2013. Retrieved 2013-07-24, from http://www.ers.usda.gov/data-products/adoption-of-genetically-engineered-crops-in-the-us/recent-trends-in-ge-adoption#.UhkNSRvVCvs.

Fernandez-Cornejo, J. and McBride, W.D. "The Adoption of Bioengineered Crops." Agricultural Economic Report 810, Economic Research Service, U.S. Department of Agriculture, 2002.

Fuglie, K., et al. "Research Investments and Market Structure in the Food Processing, Agricultural Input, and Biofuel Industries Worldwide." Economic Research Report 130, Economic Research Service, U.S. Department of Agriculture, 2011.

Gaisford, J., Hobbs, J., Kerr, W., Perdikis, N. and Plunkett, M. *The Economics of Biotechnology*. Northampton, MA: Edward Elgar, 2001.

Giorgetti, M. "Lower Bound Estimation – Quantile Regression and Simplex Method: An Application to Italian Manufacturing Sectors." *The Journal of Industrial Economics*, Vol. 51 (2003), pp. 113-120.

Goodhue, R., Rausser, G., Scotchmer, S. and Simon, L. "Biotechnology, Intellectual Property and Value Differentiation in Agriculture." CUDARE Working Paper Series 901R, Department of Agriculture and Natural Resources, University of California, Berkeley, 2002.

Graff, G., Rausser, G. and Small, A. "Agricultural Biotechnology's Complementary Intellectual Assets." *Review of Economics and Statistics*, Vol. 85 (2003), pp. 349-363.

Howell, V., Older Alguilar, D., Gleklen, J. and Letzler, K. "Competition and Innovation: A Response to the American Antitrust Institute's 'Transgenic Seed Platforms: Competition Between a Rock and a Hard Place?'." Submitted on behalf of the Monsanto Company in response to the Request for Comments by the United States Department of Agriculture and United States Department of Justice, Antitrust Division, in connection with their hearings on "Agricultural and Antitrust Enforcement Issues in Our 21st Century Economy". Cornerstone Research and Arnold & Porter, LLP, 2009.

Hubbard, K. "Out of Hand: Farmers Face the Consequences of a Consolidated Seed Industry." Farmer to Farmer Campaign on Genetic Engineering, National Family Farm Coalition, 2009.

International Science for the Acquisition of Agri-Biotech Applications (ISAAA). "Agricultural Biotechnology (A Lot More than Just GM Crops)." ISAAA SEAsiaCenter, ISAAA, 2010.

James, C. "Global Status of Commercialized Biotech/GM Crops 2010." ISAAA Brief 42-2010, International Science for the Acquisition of Agri-Biotech Applications (ISAAA), 2010.

Johnson, S.R. and Melkonyan, T.A. "Strategic Behavior and Consolidation in the Agricultural Biotechnology Industry." *American Journal of Agricultural Economics*, Vol. 85 (2003), pp. 216-233.

Kalaitzanodakes, N. and Bjornson, B. "Vertical and Horizontal Coordination in the Agro-biotechnology Industry: Evidence and Implications." *Journal of Agricultural and Applied Economics*, Vol. 29 (1997), pp. 129-139.

Latcovich, S. and Smith, H. "Pricing, Sunk Costs, and Market Structure Online: Evidence from Book Retailing." *Oxford Review of Economics*, Vol. 17 (2001), pp. 217-234.

Magnier, A., Kalaitzandonakes, N. and Miller, D. "Product Life Cycles and Innovation in the US Seed Corn Industry." *International Food and Agribusiness Management Review*, Vol. 13 (2010), pp. 17-36.

Marco, A. and Rausser, G. "The Role of Patent Rights in Mergers: Consolidation in Plant Biotechnology." *American Journal of Agricultural Economics*, Vol. 90 (2008), pp. 133-151.

Marin, P. and Siotis, G. "Innovation and Market Structure: An Empirical Evaluation of the 'Bounds Approach' in the Chemical Industry." *The Journal of Industrial Economics*, Vol. 55 (2007), pp. 93-111.

Moschini, G. "Competition Issues in the Seed Industry and the Role of Intellectual Property." *Choices*, Vol. 25 (2010).

Moss, D. "Transgenic Seed Platforms: Competition Between a Rock and a Hard Place?" White Paper, American Antitrust Institute, 2009.

Munson, C., London, S. and Lindebak, L. "Comments Regarding Competition in the Agriculture Industry." Comments submitted for the United States Department of Agriculture and United States Department of Justice Competition in the Agriculture Industry Workshops, 2010.

O'Connor, J. "Essays on Estimating Policy Effects in Science-Based Industries." Unpublished PhD Dissertation, Northwestern University, 2010.

Oehmke, J.F., Wolf, C.A. and Raper, K.C. "On Cyclical Industry Evolution in Agricultural Biotechnology R&D." *Journal of Agricultural & Food Industrial Organization*, Vol. 3 (2005), pp. 1107-1131.

Robinson, W.T. and Chiang, J. "Are Sutton's Predictions Robust?: Empirical Insights into Advertising, R&D, and Concentration." *The Journal of Industrial Economics*, Vol. 44 (1996), pp. 389-408.

Schimmelpfennig, D.E., Pray, C.E. and Brennan, M.F. "The Impact of Seed Industry Concentration on Innovation: A Study of U.S. Biotech Market Leaders." *Agricultural Economics*, Vol. 30 (2004), pp. 157-167.

Shi, G. and Chavas, J.P. "The Effects of Vertical Organization on the Pricing of Differentiated Products." *Journal of Agricultural and Resource Economics*, Vol. 36 (2011), pp. 448-464.

Shi, G., Chavas, J.P., and Stiegert, K. "An Analysis of the Pricing of Traits in the U.S. Corn Seed Market." *American Journal of Agricultural Economics*, Vol. 92 (2010), pp. 1324-1338.

Shi, G., Stiegert, K., and Chavas, J.P. "An Analysis of Bundle Pricing in Horizontal and Vertical Markets: The Case of the U.S. Cottonseed Market." Staff Paper No. 543, Department of Agricultural and Applied Economics, University of Wisconsin-Madison, 2009.

Smith, R. "Maximum Likelihood Estimation in a Class of Nonregular Cases." *Biometrika*, Vol. 72 (1985), pp. 67-90.

Smith, R. "Nonregular Regression." *Biometrika*, Vol. 81 (1994), pp. 173-183.

Stein, A.J. and Rodríguez-Cerezo, E. "What Can Data on GMO Field Release Applications in the USA Tell Us about the Commercialisation of New GM Crops?" JRC Technical Note JRC 52545, Joint Research Centre, European Commission, 2009.

Stiegert, K., Shi, G., and Chavas, J.P. "Spatial Pricing of Genetically Modified Hybrid Corn Seeds." In C. Carter, G.C. Moschini, and I. Sheldon, eds., *Genetically Modified Food and Global Welfare: Frontiers of Economics and Globalization, Vol. 10* New Milford, CT: Emerald Books, 2011.

Sutton, J. *Sunk Costs and Market Structure*. Cambridge, MA: MIT Press, 1991.

Sutton, J. *Technology and Market Structure*. Cambridge, MA: MIT Press, 1998.

Data References

BRS-APHIS-USDA. "APHIS Notification, Permit, and Petition Data." Biotechnology Regulatory Service (BRS), Animal and Plant Health Inspection Service (APHIS), United States Department of Agriculture, Washington, DC, Retrieved 2010-05-02.

CRIS-NIFA-USDA. *FY 1990 CRIS National Summary Tables*. Current Research Information System (CRIS), National Institute of Food and Agriculture (NIFA), United States Department of Agriculture (USDA), Washington, DC, 1991.

CRIS-NIFA-USDA. *FY 1991 CRIS National Summary Tables*. Current Research Information System (CRIS), National Institute of Food and Agriculture (NIFA), United States Department of Agriculture (USDA), 1992.

CRIS-NIFA-USDA. *FY 1992 CRIS National Summary Tables*. Current Research Information System (CRIS), National Institute of Food and Agriculture (NIFA), United States Department of Agriculture (USDA), 1993.

CRIS-NIFA-USDA. *FY 1993 CRIS National Summary Tables*. Current Research Information System (CRIS), National Institute of Food and Agriculture (NIFA), United States Department of Agriculture (USDA), 1994.

CRIS-NIFA-USDA. *FY 1994 CRIS National Summary Tables*. Current Research Information System (CRIS), National Institute of Food and Agriculture (NIFA), United States Department of Agriculture (USDA), 1995.

CRIS-NIFA-USDA. *FY 1995 CRIS National Summary Tables*. Current Research Information System (CRIS), National Institute of Food and Agriculture (NIFA), United States Department of Agriculture (USDA), 1996.

ERS-USDA. "Commodity Costs and Returns: Historical Costs and Returns." *Agricultural Resource Management Survey (ARMS)*, Economic Research Service (ERS), United States Department of Agriculture (USDA), Washington, DC, 2001.

ERS-USDA. "Commodity Costs and Returns: Recent Costs and Returns." *Agricultural Resource Management Survey (ARMS)*, Economic Research Service (ERS), United States Department of Agriculture (USDA), Washington, DC, 2013.

ERS-NASS-USDA. *Agricultural Chemical Usage: 1990 Field Crops Summary*. Economic Research Service (ERS), National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 1991.

ERS-NASS-USDA. *Agricultural Chemical Usage: 1991 Field Crops Summary*. Economic Research Service (ERS), National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 1992.

ERS-NASS-USDA. *Agricultural Chemical Usage: 1992 Field Crops Summary*. Economic Research Service (ERS), National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 1993.

ERS-NASS-USDA. *Agricultural Chemical Usage: 1993 Field Crops Summary*. Economic Research Service (ERS), National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 1994.

ERS-NASS-USDA. *Agricultural Chemical Usage: 1994 Field Crops Summary*. Economic Research Service (ERS), National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 1995.

ERS-NASS-USDA. *Agricultural Chemical Usage: 1995 Field Crops Summary*. Economic Research Service (ERS), National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 1996.

Fernandez-Cornejo, J. "Recent Trends in GE Adoption." Economic Research Service (ERS), US Department of Agriculture (USDA), 2013. Retrieved 2013-07-24, from http://www.ers.usda.gov/data-products/adoption-of-genetically-engineered-crops-in-the-us/recent-trends-in-ge-adoption#.UhkNSRvVCvs.

Fernandez-Cornejo, J. and McBride, W.D. "The Adoption of Bioengineered Crops." Agricultural Economic Report 810, Economic Research Service, U.S. Department of Agriculture, 2002.

Frey, K. "National Plant Breeding Study-I: Human and Financial Resources Devoted to Plant Breeding Research and Development in the United States in 1994." Special Report 98, Iowa Agriculture and Home Economics Experiment Station, Iowa State University, 1996.

NASS-USDA. *1990 Acreage Report*. National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 1990.

NASS-USDA. *1991 Acreage Report*. National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 1991.

NASS-USDA. *1992 Acreage Report*. National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 1992.

NASS-USDA. *1993 Acreage Report*. National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 1993.

NASS-USDA. *1994 Acreage Report*. National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 1994.

NASS-USDA. *1995 Acreage Report*. National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 1995.

NASS-USDA. *1996 Acreage Report*. National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 1996.

NASS-USDA. *1997 Acreage Report*. National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 1997.

NASS-USDA. *1998 Acreage Report*. National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 1998.

NASS-USDA. *1999 Acreage Report*. National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 1999.

NASS-USDA. *2000 Acreage Report*. National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 2000.

NASS-USDA. *2001 Acreage Report*. National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 2001.

NASS-USDA. *2002 Acreage Report*. National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 2002.

NASS-USDA. *2003 Acreage Report*. National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 2003.

NASS-USDA. *2004 Acreage Report*. National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 2004.

NASS-USDA. *2005 Acreage Report*. National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 2005.

NASS-USDA. *2006 Acreage Report*. National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 2006.

NASS-USDA. *2007 Acreage Report*. National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 2007.

NASS-USDA. *2008 Acreage Report*. National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 2008.

NASS-USDA. *2009 Acreage Report*. National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 2009.

NASS-USDA. *2010 Acreage Report*. National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 2010.

NASS-USDA. "United States Summary and State Data." *1992 Census of Agriculture*, Vol. 1, Chapter 2. National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA), Washington, DC, 1994.

NOAA. "Climate Indices." National Climatic Data Center (NCDC), National Oceanic and Atmospheric Administration (NOAA), US Department of Commerce. Retrieved 2011-08-28, from: https://www.ncdc.noaa.gov/cdo-web/datasets.

Rosenberg, M. "Geographic Centers of the Fifty States." About.com (web), 1999-12-06. Retrieved 2011-08-28, from: http://geography.about.com/library/weekly/aa120699a.htm.

Traxler, G., Acquaye, A.K.A., Frey, K. and Thro, A.M. "Public Sector Breeding Resources in the US: Study Results for the year 2001." Department of Agricultural Economics and Rural Sociology, Auburn University, 2005.

US Census Bureau. "State and County QuickFacts." Quickfacts.Census.Gov (web), 2011-10-19. Retrieved 2013-08-28, from: http://quickfacts.census.gov/qfd/index.html.

Online Appendix A: Sub-Market Cluster Analysis for GM Crops

The goal of cluster analysis is such that objects within a cluster (i.e., states within a regional sub-market) are "close" in terms of observable characteristics while being "far" from objects in other clusters. Thus, the objective is to define distinct, exclusive sub-markets in the agricultural seed sector by clustering states into non-overlapping partitions. We assume a "prototype-based" framework such that every state in some sub-market is more similar to some prototype state that characterizes its own sub-market relative to the prototype states that characterize other sub-markets. Therefore, we utilize a K-means approach by defining the number of sub-market clusters K for each crop type and minimizing the Euclidean distance between each state and the centroid of the corresponding cluster. For robustness, we vary the number of clusters K for each crop type and also consider alternate definitions for the distance function.

The state level data include location data (longitude and latitude measured at the state's geometric center), climate data (mean monthly temperatures, mean monthly rainfall, and mean Palmer Drought Severity Index measured by the National Oceanic and Atmospheric Administration (NOAA) from 1971-2000), and public federal funding of agricultural R&D, including USDA and CSREES (NIFA) grants, reported by the Current Research Information System (CRIS) for the fiscal years 1990-1995. The state/crop level data analyzed include farm characteristics for each crop variety (i.e., acres planted, number of farms, average farm size, number of farms participating in the retail market, total sales, and average sales per farm) that are reported by the USDA in the 1987 and 1992 US Census of Agriculture. Additional data on the application of agricultural chemicals were collected by the USDA, NASS and ERS, and reported in the *Agricultural Chemical Usage: Field Crop Summary* for the years 1990-1995.

Although there is a considerable amount of observable data on market characteristics, we encounter an issue with the degrees of freedom required for the cluster analysis when we include all available data. Specifically, the number of explanatory variables for the cluster analysis is limited to $N-K$, where N is the number of observations (i.e., states with observable characteristics) and K is the number of clusters (i.e., sub-markets). In order to reduce the problem of dimensionality in the cluster analysis, we use factor analysis, specifically principal-components factoring, to create indexes of variables that measure similar concepts (i.e., reduce monthly temperature averages to a single temperature index) and thereby reduce the number of explanatory variables.

The cluster analysis of the market for corn seed builds upon the spatial price discrimination analysis of Shi, Stiegert, and Chavas (2010). We separate the major corn production regions into "core" and "fringe" states and refine the classification of the other regions to better account for observed differences in the share of corn acres planted and proportion of acres with herbicide and/or pesticide applications prior to the introduction of GM crops. The resulting sub-markets, summarized in Table A.1, reveal that corn production is heavily concentrated in only thirteen states with Illinois and Iowa alone accounting for approximately 30% of all production.

The cluster analysis for cotton and soybean markets is slightly more problematic as fewer states farm these crops relative to corn. Regardless, the cluster analysis, along with robustness checks over the total number of clusters, reveals that the cotton and soybean markets can be reasonably divided into six sub-markets apiece, summarized in Tables A.2

and A.3, respectively. However, there are large differences in the relative size of sub-markets in cotton and soybean production as well as the regions in which production of each crop occurs. Texas accounts for over half of all planted acreage in cotton with the rest of the production primarily located in the Mississippi delta and southeast regions. Soybean production, on the other hand, primarily occurs in corn-producing regions with the significant overlap between the major corn and soybean producers.

Table A.1: Characteristics of the US Market for Corn Seed											
Sub-markets	States	Share of All Cropland Acres (1992)		Average (1990-1995) Share of Planted Corn Acres Using...					Share of Planted Acres (2010)	Share of GM Acres (2010)[1]
				Fertilizer			Herbicide	Insecticide		
		Average	Range	Nitrogen	Phosphorous	Potash	All Types	All Types		All Traits
"Core"	IL, IA	52.10%	49.25-54.82%	97.94%	81.56%	80.18%	98.23%	30.85%	29.47%	86.11%
Western "Fringe"	NE, ND, SD	22.31%	3.10-44.81%	93.32%	68.86%	29.35%	91.85%	38.45%	17.64%	92.46%
Eastern "Fringe"	KY, IN, MI, OH	38.93%	26.40-49.25%	98.02%	91.04%	85.54%	96.89%	23.08%	15.22%	79.12%
Northern "Fringe"	MN, WI	33.11%	32.01-33.64%	96.99%	89.72%	87.06%	96.34%	16.53%	12.97%	87.89%
Southern "Fringe"	KS, MO	18.08%	9.30-55.52%	98.06%	72.96%	55.75%	93.72%	33.12%	9.10%	85.46%
S. Plains/Miss. Delta	AR, LA, MS, OK, TX	4.64%	1.30-8.54%	90.24%	67.90%	34.95%	77.59%	39.99%	4.87%	-
Mid-Atlantic	DE, MD, NJ, NY, PA, VA, WV	20.60%	8.02-32.79%	93.46%	88.75%	86.78%	93.92%	22.67%	4.16%	-
Southeastern	AL, FL, GA, NC, SC, TN	17.12%	3.60-25.51%	97.79%	93.50%	93.78%	87.37%	17.42%	2.98%	-
Western	AZ, CA, CO, ID, MT, NM, OR, UT, WA, WY	3.57%	0.22-16.12%	93.61%	63.61%	32.17%	77.88%	45.79%	3.38%	-
Northeastern	CT, ME, MA, NH, RI, VT	1.70%	0.69-3.73%	91.00%	83.00%	83.00%	94.00%	23.00%	0.20%	-
US Total		23.43%	-	96.68%	82.05%	71.18%	95.57%	28.82%	-	82.00%

Sources: Authors' calculations using data collected from Census of Agriculture (1992), "Agricultural Chemical Usage: Field Crops Summary" publications by the NASS/ERS (Years: 1990-1995), and "Adoption of Genetically Engineered Crops in the US" an ERS data product (Years 2000-2010).

1: Data on adoption rates of GM varieties are unavailable for any of the states in the "Mid-Atlantic", "Southeastern", "Western", or "Northeastern" sub-markets and only Texas in the "Southern Plains/Mississippi Delta" sub-market.

Table A.2: Characteristics of the US Market for Cotton Seed											
Sub-markets	**States**	**Share of All Cropland Acres (1992)**		Average (1990-1995) Share of Planted Cotton Acres Using...					**Share of Planted Acres (2010)**	**Share of GM Acres (2010)**[2]
				Fertilizer[1]			**Herbicide**[1]	**Insecticide**[1]		
		Average	Range	Nitrogen	Phosphorous	Potash	All Types	All Types		All Traits
Texas	TX	19.96%	-	72.50%	54.67%	25.67%	93.33%	49.80%	52.42%	91.00%
Southeastern	AL, FL, GA, SC, TN	12.80%	1.74-20.52%	-	-	-	-	-	20.95%	96.11%
Mississippi Delta	AR, LA, MS	13.36%	10.96-21.72%	96.42%	54.23%	65.47%	96.57%	91.93%	10.82%	94.21%
Atlantic	NC, VA	5.87%	0.84-8.95%	-	-	-	-	-	5.87%	-
Southern Plains	KS, MO, OK	1.94%	0.01-7.11%	-	-	-	-	-	5.04%	95.27%
Southwestern	AZ, CA, NM	15.91%	5.04-47.03%	93.36%	35.74%	11.63%	79.10%	88.63%	3.19%	-
US Total		3.70%	-	83.03%	52.08%	35.93%	92.60%	68.60%	-	92.00%

Sources: Authors' calculations using data collected from Census of Agriculture (1992), "Agricultural Chemical Usage: Field Crops Summary" publications by the NASS/ERS (Years: 1990-1995), and "Adoption of Genetically Engineered Crops in the US" an ERS data product (Years 2000-2010).

1: Fertilizer, herbicide, and insecticide use on cotton crops is unavailable for Alabama, Florida, Georgia, Kansas, Missouri, New Mexico, North Carolina, Oklahoma, South Carolina, Tennessee, and Virginia

2: Data on adoption rates of GM varieties is available for only California in the "Southwestern" sub-market and only for North Carolina in the "Atlantic" sub-market. Additionally, adoption rates are unavailable for Florida, Oklahoma, and South Carolina and average GM adoption across the US is used as to approximate GM adoption in these states.

Table A.3: Characteristics of the US Market for Soybean Seed											
Sub-markets	**States**	**Share of All Cropland Acres (1992)**		Average (1990-1995) Share of Planted Soybean Acres Using...					**Share of Planted Acres (2010)**	**Share of GM Acres (2010)**[2]
				Fertilizer[1]			**Herbicide**[1]	**Insecticide**[1]		
		Average	Range	Nitrogen	Phosphorous	Potash	All Types	All Types		All Traits
Western "Core"	IA, MN, MO, WI	33.36%	6.50-95.55%	12.11%	14.93%	16.30%	97.15%	0.23%	31.53%	94.25%
Eastern "Core"	IL, IN, KY, MI, OH	36.34%	20.23-40.85%	20.38%	29.85%	38.52%	98.08%	0.93%	28.91%	89.47%
Northern Plains	KS, NE, ND, SD	9.78%	3.29-15.07%	18.20%	17.07%	8.09%	93.80%	0.65%	22.38%	95.22%
Southeastern	AL, FL, GA, NC, SC, TN	20.90%	2.04-33.50%	42.95%	56.24%	59.68%	93.46%	8.44%	9.40%	-
S. Plains/Miss. Delta	AR, LA, MS, OK, TX	13.10%	2.12-43.37%	13.70%	26.13%	26.52%	91.96%	6.20%	9.05%	95.17%
Mid-Atlantic	DE, MD, NJ, NY, PA, VA, WV	13.32%	1.36-49.30%	55.69%	55.04%	61.33%	90.42%	0.00%	2.72%	-
US Total		19.04%	-	15.92%	22.37%	25.52%	96.63%	1.60%	-	90.00%

Sources: Authors' calculations using data collected from Census of Agriculture (1992), "Agricultural Chemical Usage: Field Crops Summary" publications by the NASS/ERS (Years: 1990-1995), and "Adoption of Genetically Engineered Crops in the US" an ERS data product (Years 2000-2010).

1: Fertilizer, herbicide, and insecticide use on soybean crops is unavailable for New York and West Virginia.

2: Data on adoption rates of GM varieties is not available for either the "Southeastern" or "Mid-Atlantic" sub-markets. Additionally, adoption rates are unavailable Louisiana, Oklahoma, and Texas in the "Southern Plains/Mississippi Delta" sub-market and for Kentucky in the "Eastern 'Core'" sub-market. Average GM adoption across the US is used to approximate GM adoption in these states.

Sub-market Analysis: Shares of US Acres Planted

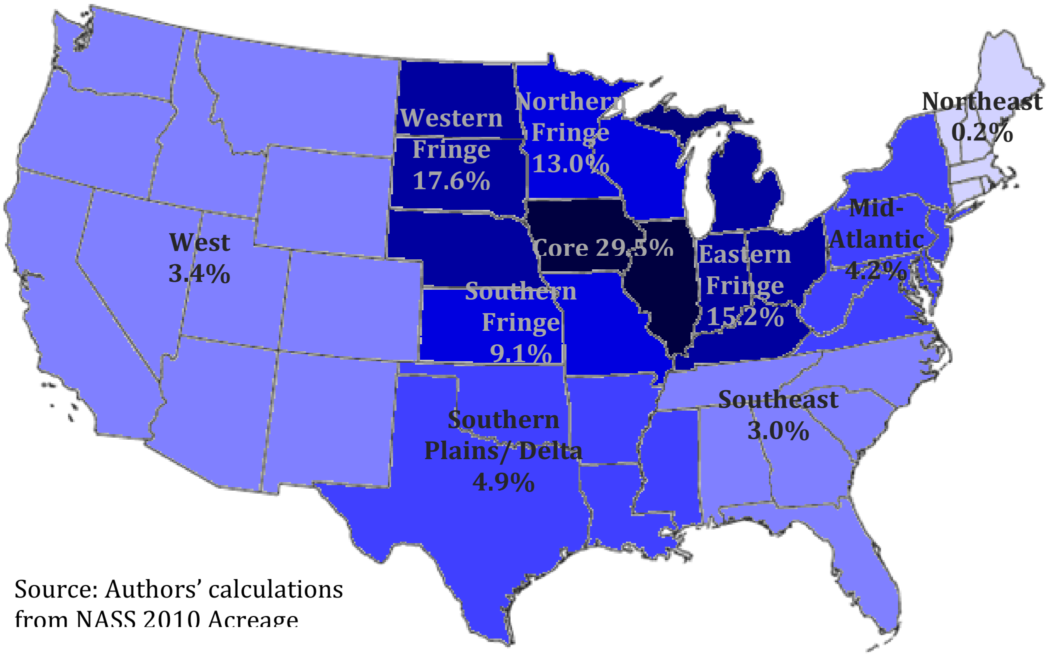


Figure A.1: 2010 Sub-market Shares of US Corn Acres Planted

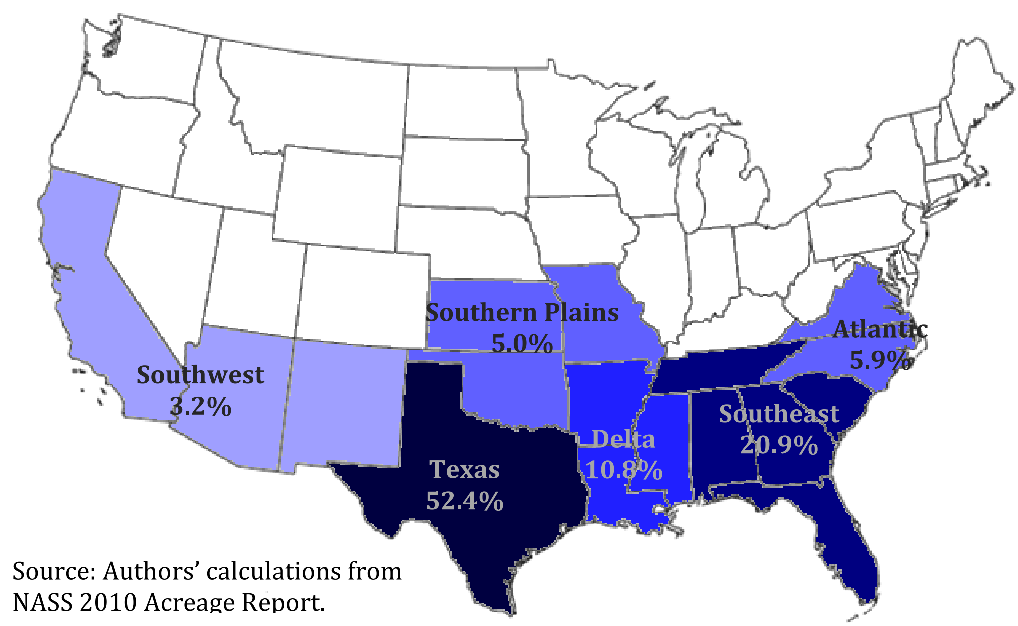
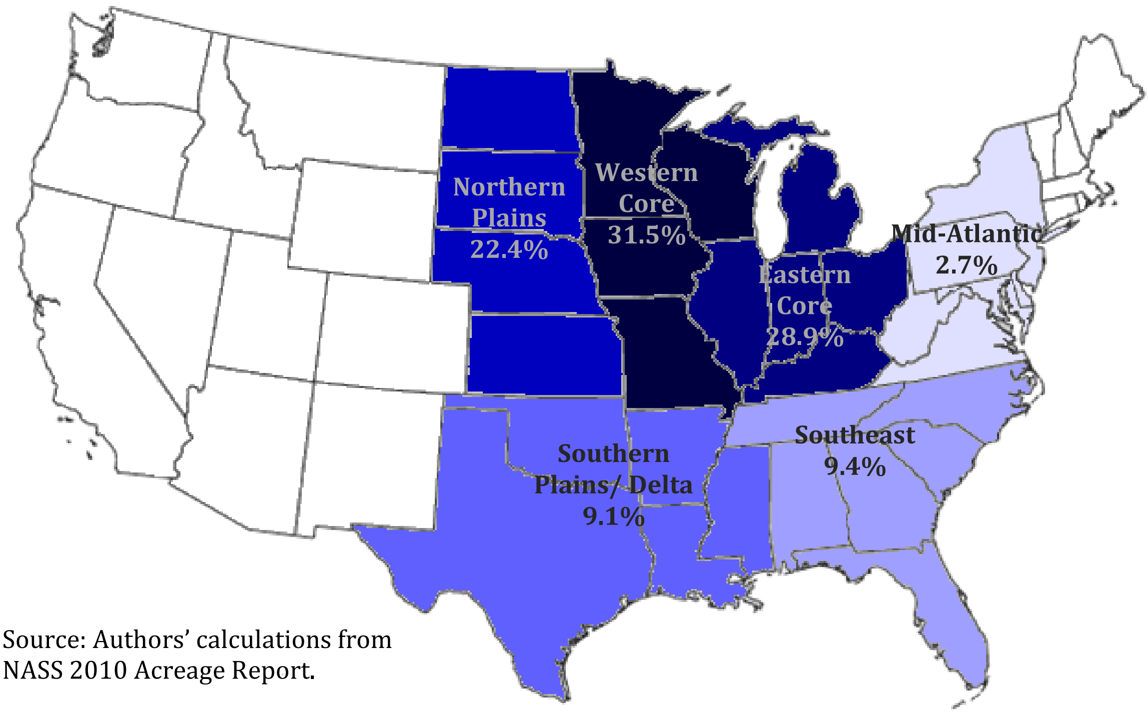


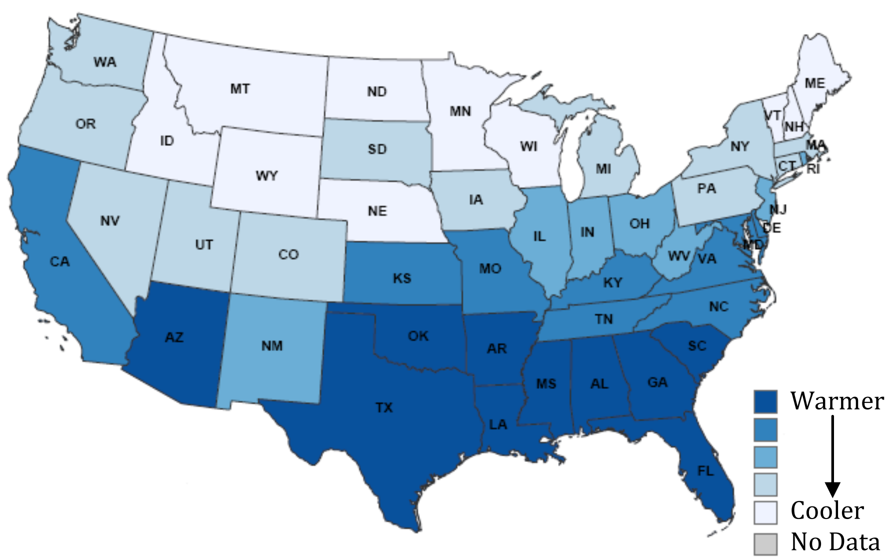
Figure A.2: 2010 Sub-market Shares of US Cotton Acres Planted



Source: Authors' calculations from NASS 2010 Acreage Report.

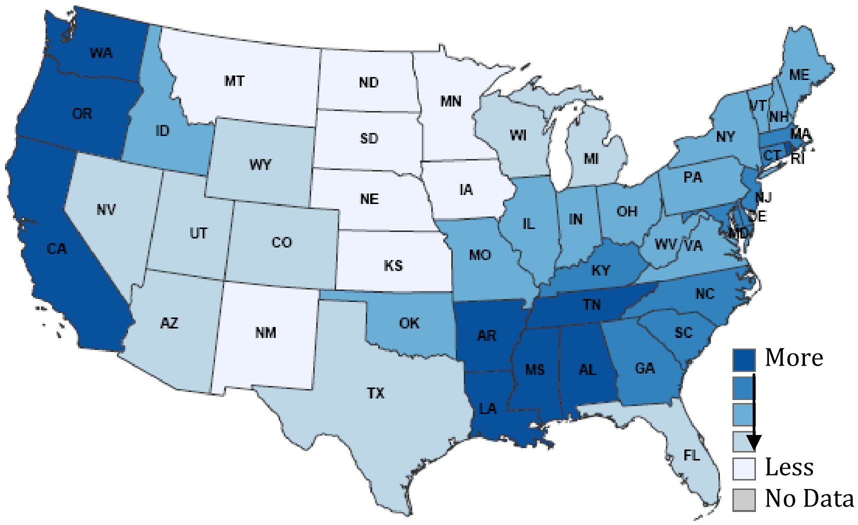
Figure A.3: 2010 Sub-market Shares of US Soybean Acres Planted

Sub-market Analysis: State-Level Climate



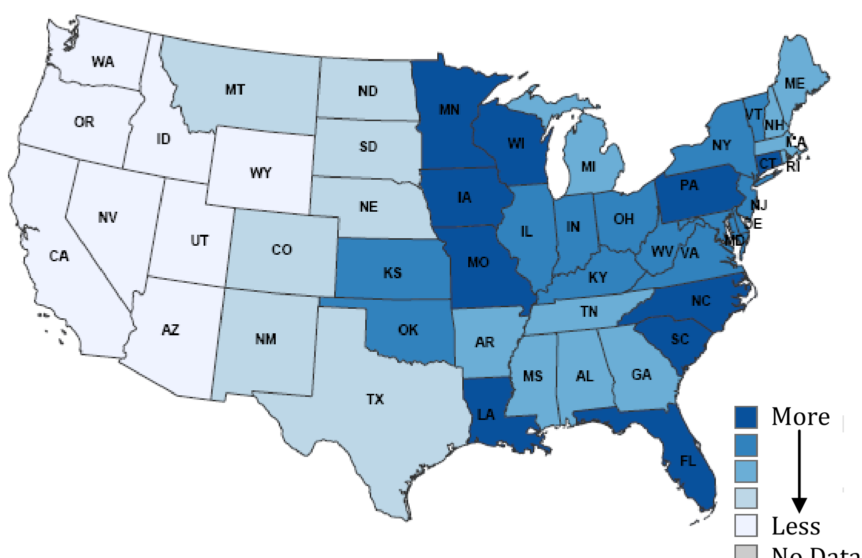
Source: Authors' calculations from NOAA data (1970-2000).

Figure A.4: Average Monthly Temperatures Factor Analysis



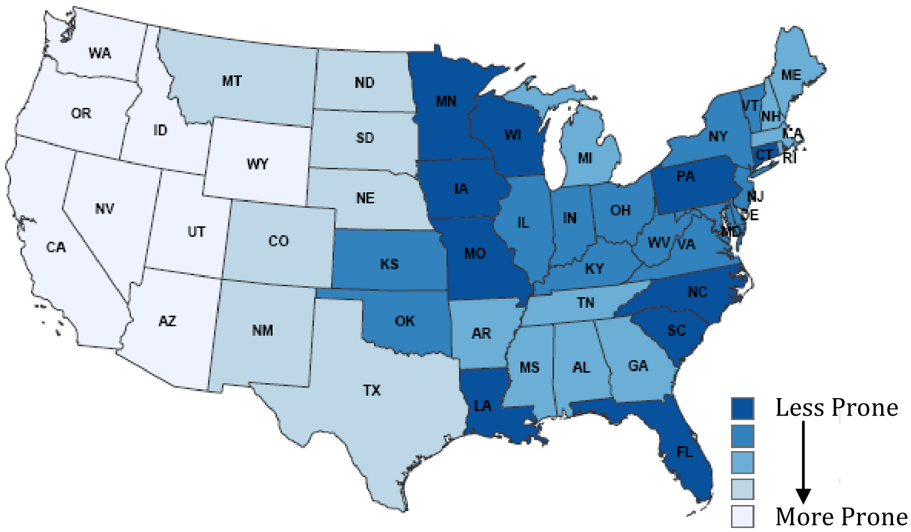
Source: Authors' calculations from NOAA data (1970-2000).

Figure A.5: Average Monthly Precipitation Factor Analysis (1)



Source: Authors' calculations from NOAA data (1970-2000).

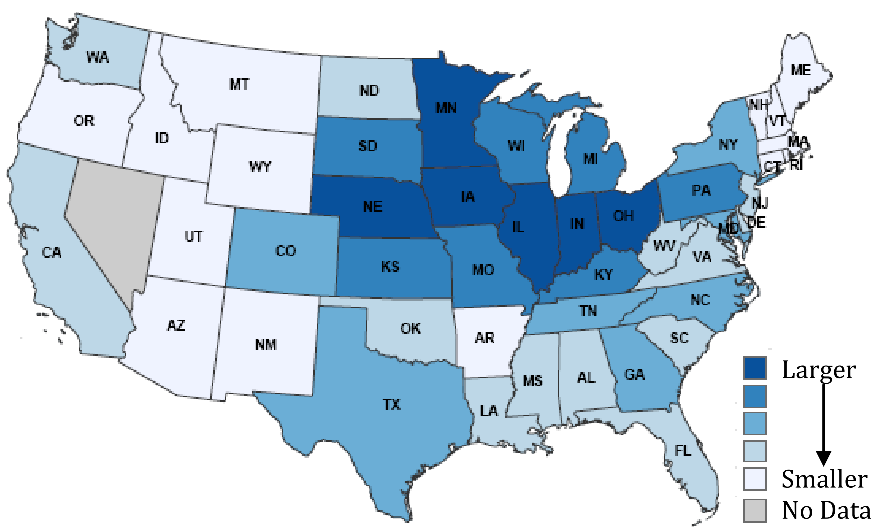
Figure A.6: Average Monthly Precipitation Factor Analysis (2)



Source: Authors' calculations from NOAA data (1970-2000).

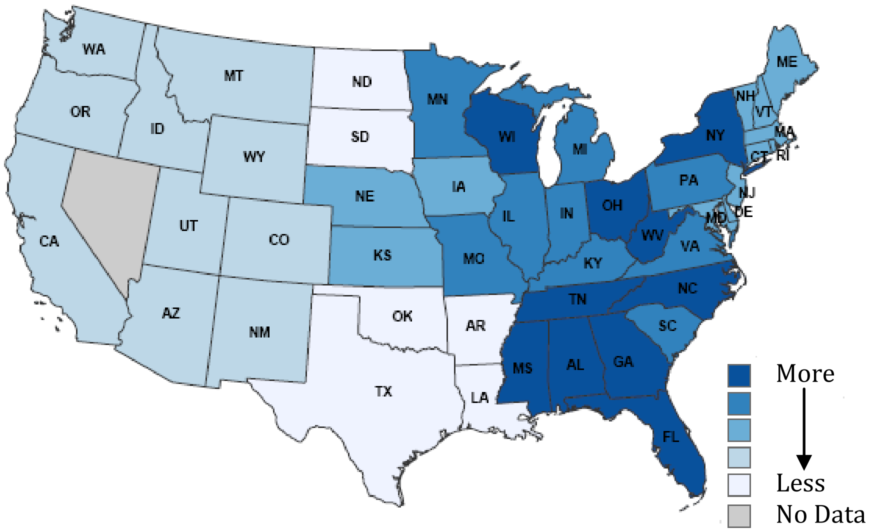
Figure A.7: Average Monthly Drought Likelihood Factor Analysis

Sub-market Analysis: Corn



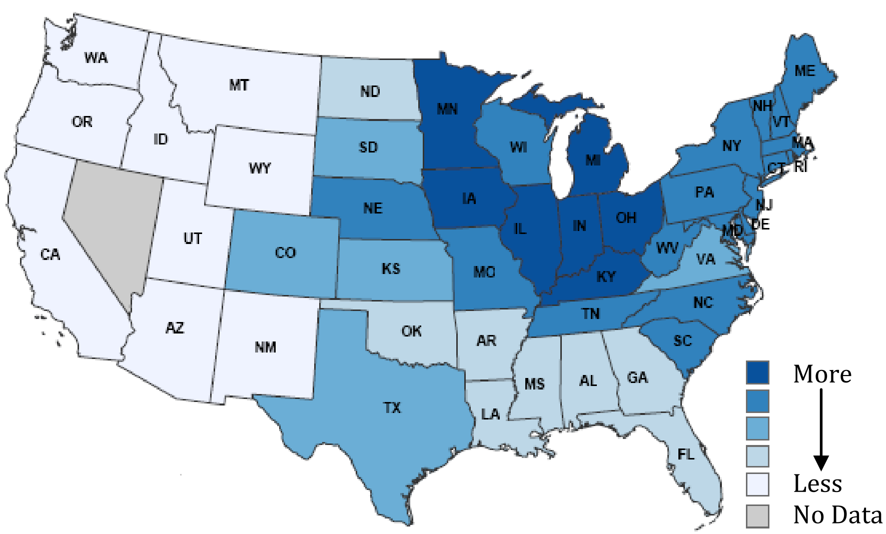
Source: Authors' calculations from Census of Agriculture (1987, 1992).

Figure A.8: Corn Seed Market Size Factor Analysis



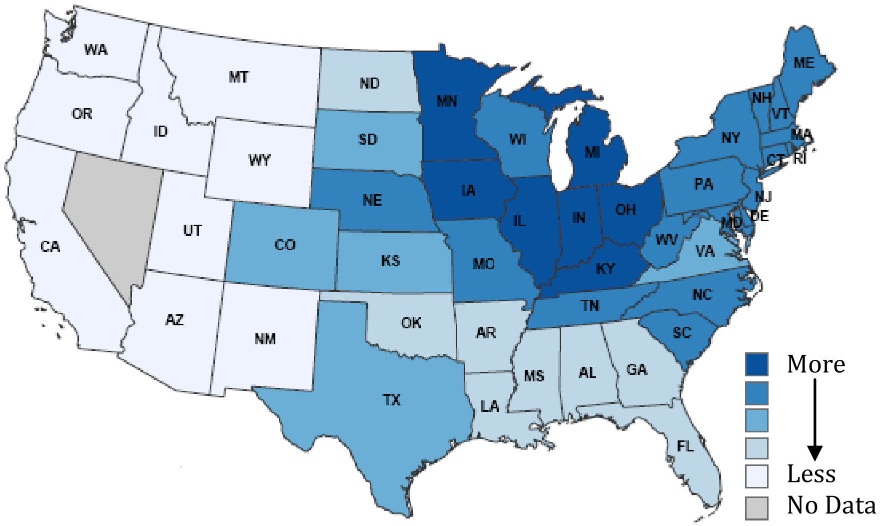
Source: Authors' calculations from Agricultural Chemical Usage (1990-1995).

Figure A.9: Percentage of Planted Corn Acres Treated with Fertilizer



Source: Authors' calculations from Agricultural Chemical Usage (1990-1995).

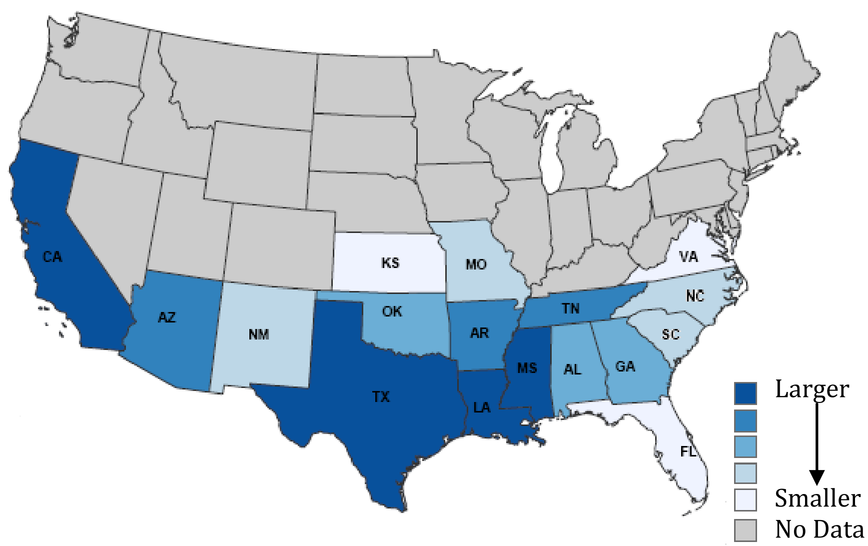
Figure A.10: Percentage of Planted Corn Acres Treated with Herbicide



Source: Authors' calculations from Agricultural Chemical Usage (1990-1995).

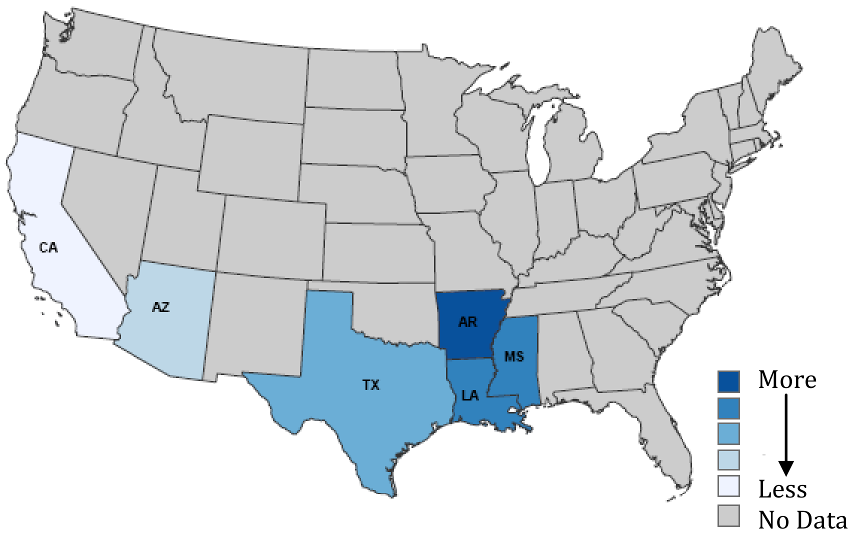
Figure A.11: Percentage of Planted Corn Acres Treated with Insecticide

Sub-market Analysis: Cotton



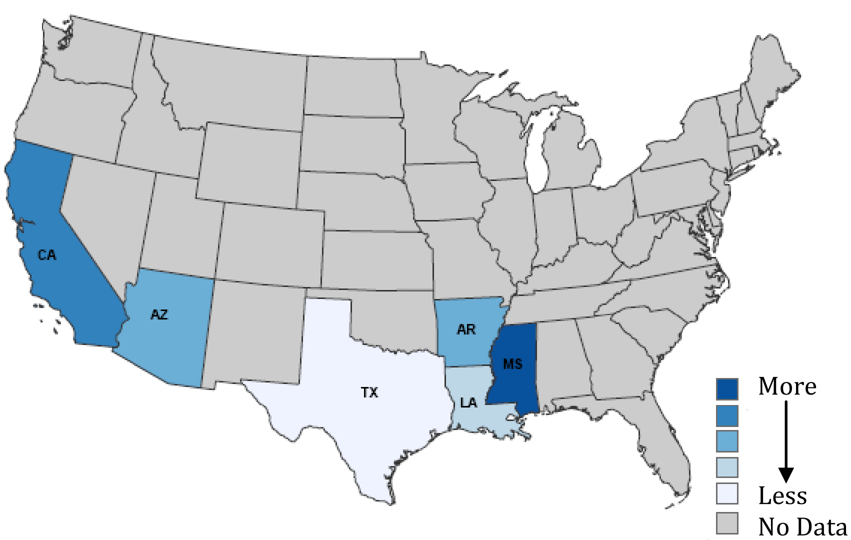
Source: Authors' calculations from Census of Agriculture (1987, 1992).

Figure A.12: Cotton Seed Market Size Factor Analysis



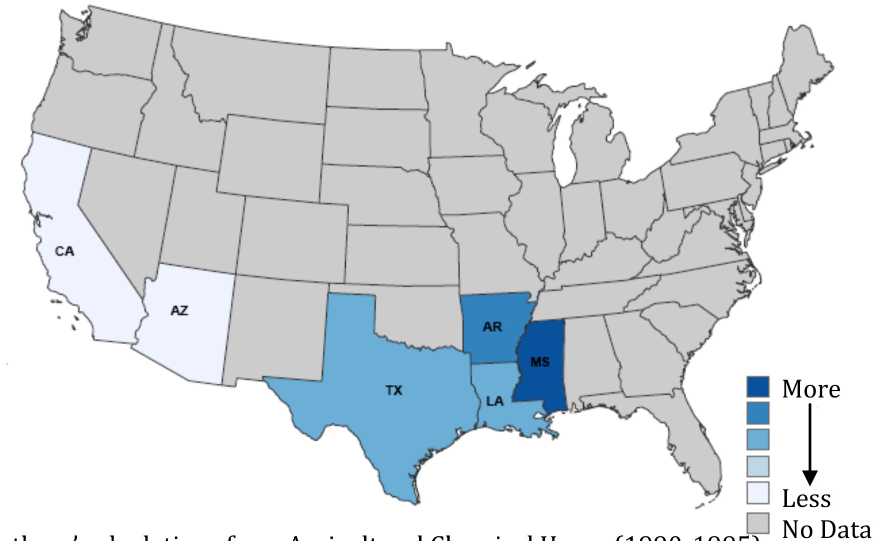
Source: Authors' calculations from Agricultural Chemical Usage (1990-1995).

Figure A.13: Percentage of Planted Cotton Acres Treated with Fertilizer (1)



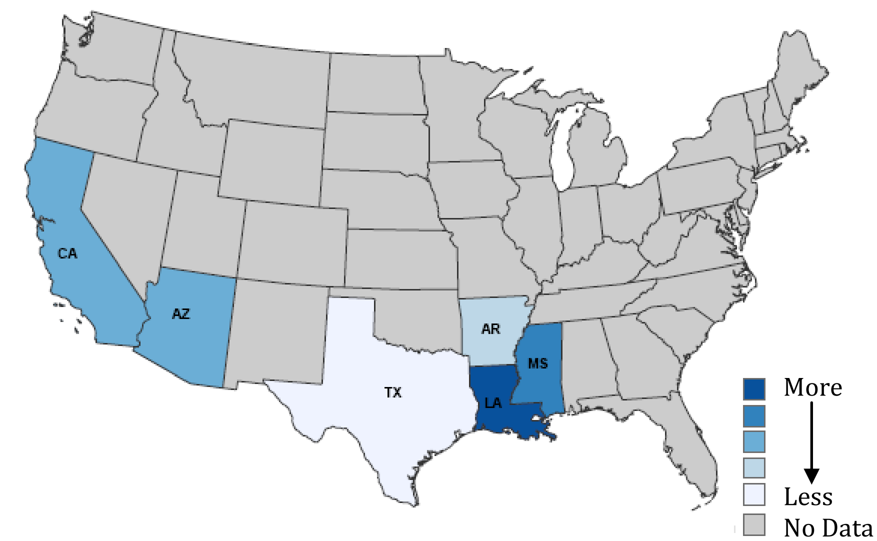
Source: Authors' calculations from Agricultural Chemical Usage (1990-1995).

Figure A.14: Percentage of Planted Cotton Acres Treated with Fertilizer (2)



Source: Authors' calculations from Agricultural Chemical Usage (1990-1995).

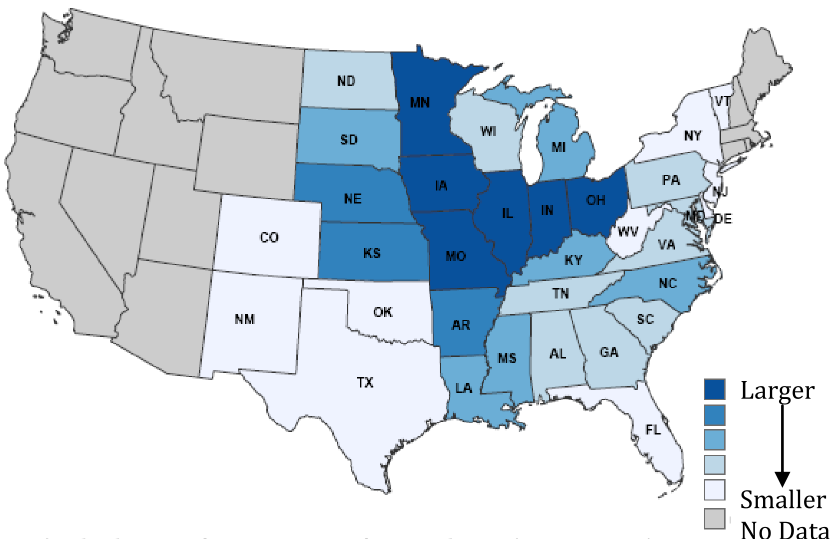
Figure A.15: Percentage of Planted Cotton Acres Treated with Herbicide



Source: Authors' calculations from Agricultural Chemical Usage (1990-1995).

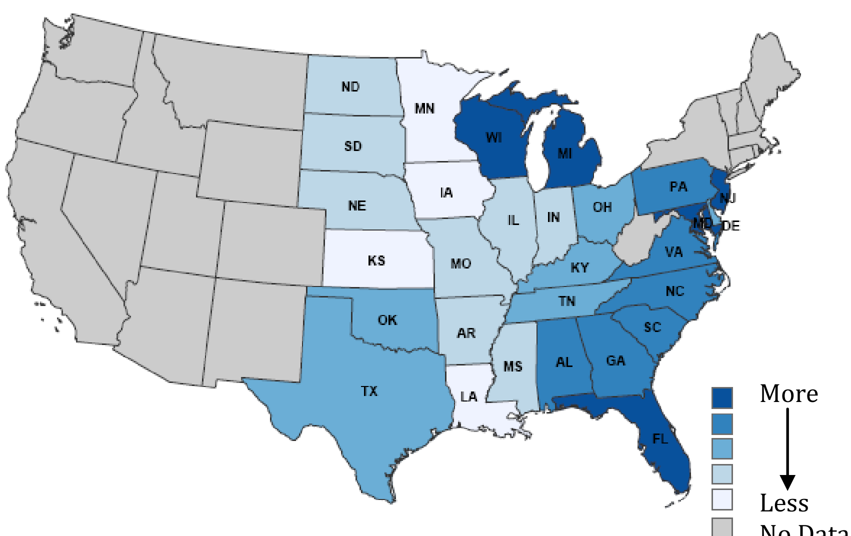
Figure A.16: Percentage of Planted Cotton Acres Treated with Insecticide

Sub-market Analysis: Soybean



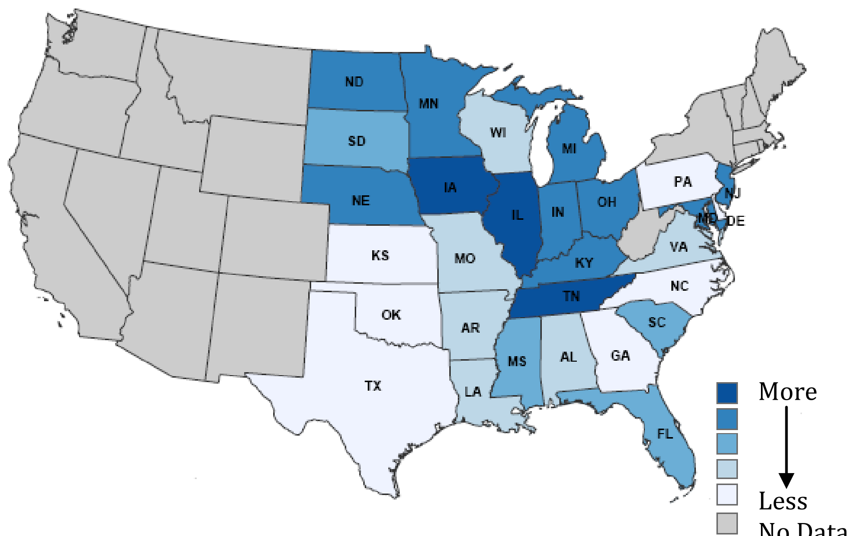
Source: Authors' calculations from Census of Agriculture (1987, 1992).

Figure A.17: Soybean Seed Market Size Factor Analysis



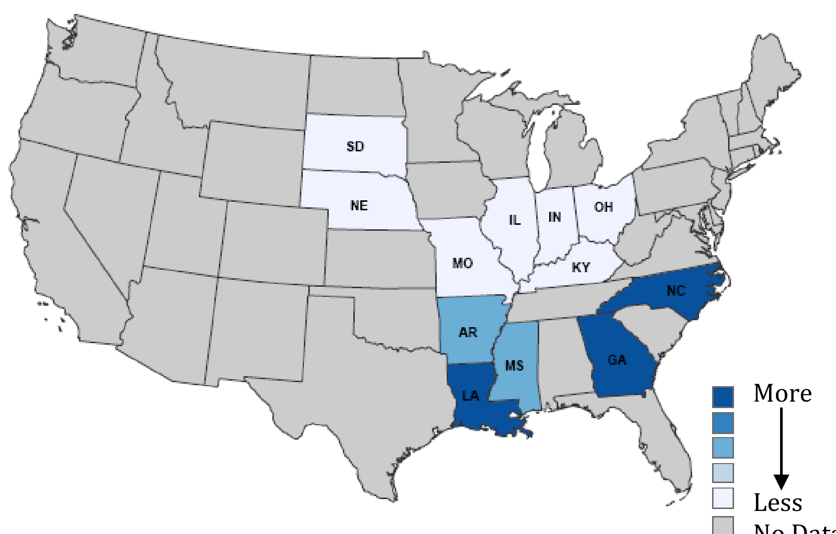
Source: Authors' calculations from Agricultural Chemical Usage (1990-1995).

Figure A.18: Percentage of Planted Soybean Acres Treated with Fertilizer



Source: Authors' calculations from Agricultural Chemical Usage (1990-1995).

Figure A.19: Percentage of Planted Soybean Acres Treated with Herbicide



Source: Authors' calculations from Agricultural Chemical Usage (1990-1995).

Figure A.20: Percentage of Planted Soybean Acres Treated with Insecticide