Endogenous R&D Investment and Market Structure: A Case Study of the Agricultural Biotechnology Industry

Benjamin C. Anderson
Department of Economics, Colgate University

Ian M. Sheldon
Department of AED Economics, The Ohio State University

April 2012

(PRELIMINARY AND INCOMPLETE – PLEASE DO NOT CITE WITHOUT PERMISSION)

Abstract
We illustrate the theoretical lower bounds to market concentration implied by an endogenous fixed cost (EFC) model with vertical and horizontal product differentiation and derive the theoretical lower bound to R&D concentration from the same model. Using data on field trial applications of genetically modified (GM) crops, we empirically 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. For the largest-sized markets in GM corn and cotton seed, single firm concentration ratios range from approximately .35 to .44 whereas three firm concentration ratios are approximately .78 to .82. The concentration ratios for GM soybean seeds are significantly lower relative to corn and cotton, despite greater levels of product homogeneity in soybeans. Moreover, adjusting for firm consolidation via mergers and acquisitions does not significantly change the lower bound estimations for the largest-sized markets in corn or cotton for either one or three firm concentration, but does increase the predicted lower bound for GM soybean seed significantly. These results imply that concentration in intellectual property in soybean varieties is differentially effected by mergers and acquisitions relative to corn and cotton varieties.

Keywords: *market structure, R&D, agricultural biotechnology*

JEL Codes: L22, Q16

***This project was supported by the Agricultural Food Research Initiative of the National Institute of Food and Agriculture, USDA, Grant #2008-35400-18704.

I. Introduction

Over the past three decades, the agricultural biotechnology sector has been characterized by rapid innovation, market consolidation, and a more exhaustive definition of property rights. Concentration has occurred in both firm and patent ownership with the six-firm concentration ratios in patents reaching approximately 50% in the U.S. and the U.K. (Harhoff, Régibeau, and Rockett, 2001) However, increased concentration has had ambiguous effects on R&D investment as the ratio of R&D expenditure to industry sales (71.4%) remains relatively large (Lavoie, 2004). Using data on field trial applications for genetically modified (GM) crops, we exploit exogenous variation in technology and market size across time and submarkets to analyze whether the agricultural biotechnology is characterized by a lower bound to concentration consistent with an endogenous fixed cost (EFC) framework.

We derive and empirically test a lower bound to R&D concentration upon the theoretical endogenous fixed cost (EFC) model of Sutton (1998). We first demonstrate the lower bounds to R&D concentration via an illustrative model and then characterize the empirical predictions from the formal model. 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.

Prior to estimating the lower bound to R&D concentration for the agricultural biotechnology industry, we examine an illustrative model of endogenous fixed costs in an industry characterized by multiple submarkets. We then derive the theoretical lower bound to R&D concentration under both exogenous and endogenous fixed costs as implied by Sutton's (1998) EFC model in order to obtain the empirical predictions for testing for a lower bound to R&D concentration. Using cluster analysis, we define regional submarkets for each GM crop type (corn, cotton, and soybean) based upon observable data on farm characteristics and crop production practices at the state level.

Ultimately, this leads to a test of the hypothesis that the agricultural biotechnology sector is characterized by an EFC model through the examination of data on field trial applications for GM crop release. The Animal and Plant Health Inspection Service (APHIS) provides data on permit, notification, and petition applications for the importation, interstate movement, and release of genetically-modified organisms in the US for the years 1985-2010. By classifying the permit data according to type, we obtain estimates of concentration in intellectual property within distinct submarkets as a measure of (intermediate) R&D concentration. We exploit variation along two dimensions: (i) geographically as adoption rates for GM crop varieties varies by state and agricultural region; and (ii) intertemporally as adoption rates for GM crops has been steadily increasing over time. Moreover, the strengthening of property rights over GM crops over the past 20 years and increased incentives for farmers to plant corn seed, relative to soybean seed, arising from the subsidies to ethanol production serve as natural experiments and provide sources of exogenous variation in the market. 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.

Results from the empirical estimations support the hypothesis that the agricultural biotechnology sector is characterized by endogenous fixed costs to R&D with the largest effects within the GM corn and cotton seed markets. However, the estimation results also indicate that within the soybean seed markets, firm merger and acquisition activity has significantly increased the observed levels of concentration in intellectual property. These results jointly reveal a difficulty associated with examinations of the agricultural biotechnology sector; namely, the nature of technology competition implies a level of concentration is to be expected, but the level of merger and acquisition activity remains an important determinate into examinations of concentration in intellectual property.

Rapid technological innovation and observed firm consolidation has led to several empirical examinations of market structure in the agricultural biotechnology industry. Fulton and Giannakas (2001) find that the agricultural biotechnology sector has undergone a restructuring in the form of both horizontal and vertical integration over the past ten years. 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 and substitutes across firms; and (iii) property rights governing plant and seed varieties that have become more clearly defined since the 1970s. This proposed research extends the stylized facts for the agricultural biotechnology industry by identifying the relevance of sunk costs investments in R&D in shaping the observed concentration and distribution of firms. As Sheldon (2008) identifies, the presence of endogenous sunk costs in R&D expenditures,

[1] In regards to economies of scale and/or economies of scope in agricultural biotechnology, 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 any conclusive results concerning economies of scale or correlation between the size of firms and the size of R&D in agricultural biotechnology.

high levels of market concentration, and high levels of R&D intensity in the agricultural biotechnology make this sector a likely candidate to be well-described by an EFC-type model such as that proposed by Sutton (1998).

In estimating an EFC-type model, this analysis extends the previous work by considering a more general framework in which concentration and innovation are jointly determined. Previously, Schimmelpfennig, Pray, and Brennan (2004) tested Schumpeterian hypotheses regarding the levels of industry concentration and innovation in biotechnology and found a negative and endogenous relationship between measures of industry concentration and R&D intensity. Additional stylized examinations of the agricultural biotechnology industry have identified an endogenous, cyclical relationship between industry concentration and R&D intensity (Oehmke, Wolf, and Raper, 2005) and categorized the endogenous relationship between firm innovation strategies, including the role of complementary intellectual assets, and industry consolidation characteristics (Kalaitzandonakes and Bjornson, 1997). As a more general model, this analysis embeds previous results that observe an endogenous relationship between R&D investments and industry concentration. Moreover, we incorporate exogenous variations in total market size for each crop type as well as technological innovations, including the development of second- and third-generation GM crops, and changes in consumer preferences over the relevant time frame to provide a richer analysis of industry configurations. 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.

A related vein of research has focused upon the significant levels of merger and acquisition activity that have historically been observed in the agricultural biotechnology industry. The explanations behind the high levels of activity 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). This analysis extends previous examinations into merger and acquisition activity in agricultural biotechnology in estimating whether this 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.

The EFC model employed in this framework has been utilized to empirically examine a variety of other industries including chemical manufacturing (Marin and Siotis, 2007), supermarkets (Ellickson, 2007), banking (Dick, 2007), newspapers and restaurants (Berry and Waldfogel, 2003), and online book retailers (Latcovich and Smith, 2001). These previous 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, 2003), or advertising (Latcovich and Smith, 2001). 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. 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. Moreover, 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.

In light of the recent Justice Department announcement regarding its investigations into anticompetitive practices in agriculture[2], this analysis is of interest to both regulators and policymakers concerned with the observed high levels of concentration in agricultural biotechnology. Specifically, if the agricultural biotechnology sector is characterized by endogenous fixed costs, the high levels of concentration, accompanied with high levels of innovative activity, are a natural outcome of technology competition and are not evidence of collusion among firms. However, the significant shift in the observed patterns of R&D concentration in cotton and soybean seed upon accounting for merger and acquisition activity imply that industry consolidation has increased concentration of intellectual property to levels greater than what is predicted under endogenous fixed costs alone.

---

[2] Neuman, W. 2010. "Justice Dept. Tells Farmers It Will Press Agricultural Industry on Antitrust." *The New York Times*, March 13, pp. 7.

II. What is Agricultural Biotechnology?

Prior to examining market structure and innovation in the agricultural biotechnology sector, it is important to clearly define what we mean when we use the term "agricultural biotechnology". Gaisford, et al. (2001) define biotechnology, in general terms, 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." For our purposes, we are interested in firms that develop genetically-modified organisms (GMOs) for commercialization purposes within agriculture and restrict ourselves primarily to discussion concerning genetically-modified (GM), or genetically-engineered (GE), crops.

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 (Lavoie, 2004). These advances had two key implications for agricultural seed manufacturers and plant and animal scientists. 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.

GM crops are typically assigned into three broad classifications, termed "generations", depending upon the traits that they display and who benefits from these technological advancements (i.e. farmers, consumers, or other firms). The first generation consists of crops that display cost- and/or risk-reducing traits that primarily benefit the farmers, but which also may have important environmental and consumer impacts via decreased application of agricultural chemicals. Specific examples of first generation crops include herbicide tolerant varieties (i.e. Roundup Ready® crops), insect resistant varieties (i.e. Bt crops), or crop types that are particularly tolerant to environmental stresses including drought or flood (Fernandez-Cornejo and Caswell, 2006). Second generation crops, which are largely still in development, consist of crops whose final products will deliver some additional value-added benefits directly to consumers. Products derived from second generation crops might offer increased nutritional content or other characteristics that directly benefit the health of end consumers. The third generation classification captures biotechnology crops developed for pharmaceuticals, industrial inputs (i.e. specialized oils or fibers), or bio-based fuels. We focus almost exclusively upon crops within the "first generation" classification as these constitute the majority of all currently commercialized GM crops. However, our analysis is applicable to the biotechnology industry in a general sense to the extent that we identify how the industry has evolved in the past with implications for how market structure and innovation will evolve as subsequent generations of biotechnology are introduced.

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

An Illustrative Model

We adapt the theoretical endogenous fixed cost model of market structure and sunk R&D investments developed by Sutton (1998) and empirically estimate the lower bounds to R&D concentration in agricultural biotechnology. 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). We illustrate that the characterization of a market into horizontally differentiated submarkets does not change the theoretical predictions for the lower bound to concentration in the largest submarket under endogenous fixed costs. Subsequently, we derive the theoretical lower bound to R&D concentration for endogenous and exogenous fixed cost industries and specify the empirically testable hypotheses.

The specification of the empirical model 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 submarkets. This assumption corresponds with the empirical findings of Stiegert, Shi, and Chavas (2011) of spatial price differentiation in GM corn. Secondly, we assume that farmers value higher quality products such that a firm competes within each submarket 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 submarkets.

In order to derive the empirical predictions for the lower bound to R&D concentration, we adapt the illustrative model developed by Sutton (1991). We assume that within a regional submarket m, there are S_m identical farmers that have a quality-indexed demand function such that:

$$\mathrm{U}_m = (uq)^{\gamma_m} z^{1-\gamma_m}, \tag{1}$$

where z is some "outside" composite good (i.e. fertilizer, machinery, etc.) which is set as numeraire, q is the quantity of the "quality" good (i.e. seeds), u is the quality level associated with good q and preferences are captured by the parameter γ_m. We assume a level of quality $u \geq 1$ such that $u = 1$ corresponds to a minimum level of quality in the market and all farmers prefer higher quality for a given set of prices. The farmer in submarket m maximizes across all quality goods such that:

$$\begin{aligned} &\max_{\{q_{im}\},z} \; (u_{im}q_{im})^{\gamma_m} z^{1-\gamma_m} \\ &s.t. \quad p_{im}q_{im} + p_z z \leq M_m \end{aligned}, \tag{2}$$

where M_m is the total income for the farmer in submarket m. Solving reveals that, independent of equilibrium prices or qualities, the farmer will spend a fraction γ_m of her total income upon the quality good.

We consider a three stage game consisting of: (i) a market entry decision into some submarket m; (ii) technology market competition in which firms make fixed R&D

investments in product quality; and (iii) product market competition in quantities. In the second stage, given decisions to enter in the first stage, firms choose the levels of quality $u \geq 1$ they offer by making deterministic fixed (sunk) R&D investments within each submarket. In the final stage, firms engage in Cournot competition over quantities in the product market with the set of product quality levels $\{u_i\}$ taken as given. The farmer thus chooses the good that maximizes the quality-price ratio u_i/p_i such that all firms that have positive sales in equilibrium have proportionate quality-price ratios (i.e. $u_i/p_i = u_j/p_j, \forall i, j$).

In order to derive the profit function for firms, consider the case in which all firms in a submarket are symmetric (i.e. $u_{im} = \bar{u}_m, \forall i$). It must be the case that the equilibrium level of prices $\bar{p}_m$ equals the share of expenditure over total industry output $\sum_i \bar{q}_{im}$ such that:

$$\bar{p}_m = \frac{S_m \gamma_m M_m}{\sum_i \bar{q}_{im}}. \tag{3}$$

Now suppose a single firm i deviates from the symmetric equilibrium by offering a quality level u_{im} such that it faces a price p_{im} equal to:

$$p_{im} = \frac{u_{im}}{\bar{u}_m} \bar{p}_m. \tag{4}$$

It follows that the equilibrium price $\bar{p}_m$ faced by all other firms can be expressed as:

$$\bar{p}_m = \frac{S_m \gamma_m M_m}{\sum_{j \neq i} \bar{q}_{jm} + \left(\frac{u_{im}}{\bar{u}_m}\right) q_{im}}, \tag{5}$$

where $\sum_{j \neq i} \bar{q}_{jm}$ is total industry output net the output of the deviating firm and q_{im} is the deviating firm output. Assuming that firms face a constant marginal cost c independent of the level of quality offered, the profit functions for the deviating firm i and any non-deviating firm j in submarket m can thus be expressed, respectively, as:

$$\pi_{im} = (p_{im} - c)q_{im} = \left(\frac{u_{im}}{\bar{u}_m} \cdot \frac{S_m \gamma_m M_m}{\sum_{j \neq i} \bar{q}_{jm} + \frac{u_{im}}{\bar{u}_m} q_{im}} - c\right) \cdot q_{im} \tag{6}$$

and

$$\bar{\pi}_{jm} = (\bar{p}_m - c)\bar{q}_{jm} = \left(\frac{S_m \gamma_m M_m}{\sum_{j \neq i} \bar{q}_{jm} + \frac{u_{im}}{\bar{u}_m} q_{im}} - c\right) \cdot \bar{q}_{jm}. \tag{7}$$

Differentiating and solving for the equilibrium levels of quantity yields the following expressions for the quantity produced for the deviating firm i and $(N_m - 1)$ non-deviating firms as a function of quality levels u_{im} and $\bar{u}_m$ such that:

$$q_{im}^* = \left[(N_m - 2) - \left(\frac{\bar{u}_m}{u_{im}}\right)(N_m - 1)\right] \cdot \bar{q}_m \tag{8}$$

and

$$\bar{q}_m = \frac{S_m \gamma_m M_m}{c} \left[\frac{\left(\frac{u_{im}}{\bar{u}_m}\right)(N_m - 1)}{\left[1 + \left(\frac{u_{im}}{\bar{u}_m}\right)(N_m - 1)\right]^2}\right]. \tag{9}$$

Substituting the expressions for the equilibrium levels of quantity (8) and (9) into the expressions for prices (4) and (5), we derive the equilibrium prices for the deviating firm i and $(N_m - 1)$ non-deviating firms such that:

$$p_{im} = c\left[\left(\frac{u_{im}}{\bar{u}_m}\right) + \frac{1}{(N_m - 1)}\right] \tag{10}$$

and

$$\bar{p}_m = c\left[1 + \frac{1}{\left(\frac{u_{im}}{\bar{u}_m}\right)(N_m - 1)}\right]. \tag{11}$$

Thus, the net final-stage profit for the deviating firm in submarket m can be expressed as:

$$\pi(u_{im}|\bar{u}_m, N_m) = S_m\gamma_m M_m\left[1 - \frac{1}{\left(\frac{u_{im}}{\bar{u}_m}\right) + \frac{1}{(N_m - 1)}}\right]^2. \tag{12}$$

This profit function allows for the examination of the case where all firms enter with symmetric quality (i.e. $u_{im} = \bar{u}_m$) and earn final-stage profits independent of quality (i.e. $\pi_m = \frac{S_m\gamma_m M_m}{N_m^2}$) as well as the case in which firms make fixed (sunk) R&D investments in quality. Letting the total market size (i.e. $S_m\gamma_m M_m$) in submarket m be Γ_m and summing across all submarkets in which firm i is active (i.e $m \in I$), firm i's total profit Π_i can be expressed as:

$$\Pi_i = \sum_{m\in I}\Gamma_m\left[1 - \frac{1}{\left(\frac{u_{im}}{\bar{u}_m}\right) + \frac{1}{(N_m - 1)}}\right]^2. \tag{13}$$

Sutton (1998) proposes a specification for product quality given the possibility of economies of scope across R&D trajectories. If agricultural biotechnology firms develop seed varieties that share attribute traits in adjoining geographic submarkets, then such a specification could capture technology spillovers between submarkets. Therefore, the quality level offered by some firm i in submarket m can be expressed as a function of the competencies that the firm achieves along all research trajectories such that:

$$u_{im} = v_{im} + \sigma\sum_{n\in I\backslash m} v_{in}, \tag{14}$$

where v_{im} is the competency that firm i achieves in submarket m and $\sigma \in [0,1]$ is a measure of economies of scope across competencies.

We assume a R&D cost function that consists of a minimum setup cost F_0 associated with entry in each submarket and a variable component that is increasing in the level of competency v. Thus for a given geographic submarket m (i.e. research trajectory), firm i chooses a competency v_{im} and incurs a sunk R&D cost $F(v_{im})$ equal to:

$$F(v_{im}) = F_0 v_{im}^{\beta}, \tag{15}$$

where β is the elasticity of the fixed cost schedule. We assume $\beta > 2$ such that R&D investment rises with quality at least as fast as profits for a given increase in quality. We obtain an expression for firm i's total R&D investment F_i by summing across geographic submarkets (i.e. research trajectories) such that:

$$F_i = \sum_{m \in I} F_0\left(v_{im}^{\beta} - 1\right) + n_i F_0, \tag{16}$$

where n_i corresponds to the total number of submarkets that firm i enters.

Given the expressions for firm profit (13) and firm R&D costs (16), the firm's payoff function (i.e. the profit function net of fixed R&D investments) for the second stage quality choice decision can be written as:

$$\Pi_i - F_i = \sum_{m \in I} \left\{ \Gamma_m \left[1 - \frac{1}{\left(\frac{u_{im}}{\bar{u}_m}\right) + \frac{1}{(N_m - 1)}} \right]^2 - F_0\left(v_{im}^{\beta} - 1\right) \right\} + n_i F_0, \tag{17}$$

where firms take the number of entrants N_m in each submarket as given from the first-stage entry decision.

We assume that each submarket in which firms enter can support at least a single firm producing minimum quality such that:

$$\Gamma_m \geq F_0. \tag{18}$$

Given that the assumption on the size of the market and minimum setup cost for entry holds, we identify two possible symmetric equilibrium outcomes by solving the quality choice condition in the second stage. The first case corresponds with a symmetric equilibrium in which all firms active firms in submarket m enter with minimum quality (i.e. $v_{im} = 1, \forall i$) and incur only the minimum setup cost F_0 such that:

$$\left.\frac{\partial \pi_{im}}{\partial v_{im}}\right|_{v_{im}=\bar{v}_m=1} \leq \left.\frac{dF}{dv_{im}}\right|_{v_{im}=\bar{v}_m=1}. \tag{19}$$

Condition (19) is equivalent to the case of exogenous fixed costs in which all firms enter with minimum quality (i.e. $v_{im} = \bar{v}_m = 1, \forall i$).

We define symmetric free entry conditions for each submarket m such that firms enter in the first stage until additional entrants are unable to recoup their fixed R&D investments in the submarket such that:

$$\forall m \quad \pi(\bar{u}_m | \bar{u}_m, N_m) = F\left(\bar{v}_m^{\beta}\right). \tag{20}$$

We now derive the number of firms N_m^{EX} that enter in a symmetric equilibrium and incur only the R&D setup cost F_0 by investing in the minimum competency level ($\bar{v}_m = 1$) in submarket m. Therefore, the free entry condition (20) under exogenous fixed costs by can be expressed as:

$$\Gamma_m \left[1 - \frac{1}{1 + \frac{1}{(N_m - 1)}} \right]^2 = F_0. \tag{21}$$

Solving condition (21) explicitly for the equilibrium number of firms N_m^{EX} yields:

$$N_m^{EX} = \sqrt{\frac{\Gamma_m}{F_0}}, \tag{22}$$

which depends upon the market size of submarket m (i.e. the number of consumers S_m, the proportion of income spent on the "quality" good γ_m, and the income of consumers M_m) and the minimum setup cost F_0.

If condition (19) does not hold, then the quality choice condition for a symmetric equilibrium in which all firms enter with quality greater than the minimum (i.e. $v_{im} = \bar{v}_m > 1, \forall i$) can be expressed as:

$$\left.\frac{\partial \pi_{im}}{\partial v_{im}}\right|_{v_{im}=\bar{v}_m>1} = \left.\frac{dF}{dv_{im}}\right|_{v_{im}=\bar{v}_m>1}. \tag{23}$$

Condition (23) is the case in which firms make quality-enhancing investments in R&D such that fixed costs are endogenous. Expressing condition (23) explicitly yields:

$$\frac{2\Gamma_m}{\bar{u}_m} \cdot \left[\frac{(N_m - 1)^2}{N_m^3} \right] + 2\sigma \sum_{l \in I \setminus m} \frac{\Gamma_l}{\bar{u}_l} \cdot \left[\frac{(N_l - 1)^2}{N_l^3} \right] = \beta F_0 \bar{v}_m^{\beta - 1}. \tag{24}$$

Adding and subtracting $\frac{2\sigma\Gamma_m}{\bar{u}_m} \cdot \left[\frac{(N_m-1)^2}{N_m^3}\right]$, we can express condition (24) as:

$$(1 - \sigma) \frac{\Gamma_m}{\bar{u}_m} \cdot \left[\frac{(N_m - 1)^2}{N_m^3} \right] + \sigma \sum_{l \in I} \frac{\Gamma_l}{\bar{u}_l} \cdot \left[\frac{(N_l - 1)^2}{N_l^3} \right] = \frac{\beta}{2} F_0 \bar{v}_m^{\beta - 1}.$$

Substituting equation (14) for $\bar{u}_m$ and adding and subtracting $\sigma \bar{v}_m$ yields:

$$(1-\sigma)\left[(1-\sigma)\bar{v}_m+\sigma\sum_{n\in I}\bar{v}_n\right]^{-1}\left[\frac{\Gamma_m(N_m-1)^2}{N_m^3}\right]$$
$$+\sigma\sum_{l\in I}\left[(1-\sigma)\bar{v}_l+\sigma\sum_{n\in I}\bar{v}_n\right]^{-1}\left[\frac{\Gamma_l(N_l-1)^2}{N_l^3}\right]=\frac{\beta}{2}F_0\bar{v}_m^{\beta-1}.$$

Letting $v_i=\sum_{n\in I}\bar{v}_n$ and $\phi_m=\left[\frac{\Gamma_m(N_m-1)^2}{N_m^3}\right]$ and multiplying both sides by $\bar{v}_m$ yields:

$$(1-\sigma)\bar{v}_m[(1-\sigma)\bar{v}_m+\sigma v_i]^{-1}\phi_m$$
$$+\sigma\bar{v}_m\sum_{l\in I}[(1-\sigma)\bar{v}_l+\sigma v_i]^{-1}\phi_l=\frac{\beta}{2}F_0\bar{v}_m^{\beta}. \tag{25}$$

Summing equation (25) across all $m\in I$ such that:

$$(1-\sigma)\sum_{m\in I}\bar{v}_m[(1-\sigma)\bar{v}_m+\sigma v_i]^{-1}\phi_m$$
$$+\sigma\sum_{l\in I}[(1-\sigma)\bar{v}_l+\sigma v_i]^{-1}\phi_l\sum_{m\in I}\bar{v}_m=\frac{\beta}{2}F_0\bar{v}_m^{\beta}. \tag{26}$$

Since we are summing across all submarkets in which firm i is active, equation (26) can be simplified such that:

$$\sum_{m\in I}[(1-\sigma)\bar{v}_m[(1-\sigma)\bar{v}_m+\sigma v_i]^{-1}\phi_m+\sigma v_i[(1-\sigma)\bar{v}_m+\sigma v_i]^{-1}\phi_m]=\frac{\beta}{2}F_0\sum_{m\in I}\bar{v}_m^{\beta}.$$

Collecting terms, simplifying, and substituting for ϕ_m yields:

$$\sum_{m\in I}\left[\frac{\Gamma_m(N_m-1)^2}{N_m^3}\right]=\frac{\beta}{2}F_0\sum_{m\in I}\bar{v}_m^{\beta}. \tag{27}$$

We state the free entry condition, as characterized by equation (20), when firms enter symmetrically with competency $\bar{v}_m>1$ in submarket m explicitly as:

$$\frac{\Gamma_m}{N_m^2}=F_0\bar{v}_m^{\beta}, \tag{28}$$

such that summing expression (28) across all submarkets in which some firm i is active (i.e. $m \in I$) yields:

$$\sum_{m \in I} \frac{\Gamma_m}{N_m^2} = F_0 \sum_{m \in I} \bar{v}_m^{\beta}. \tag{29}$$

Dividing both sides of the quality choice condition (27) by both sides of the free entry condition (29) yields an expression for the equilibrium number of firms across submarkets such that:

$$\frac{\sum_{m \in I} \left[\frac{\Gamma_m (N_m - 1)^2}{N_m^3}\right]}{\sum_{m \in I} \frac{\Gamma_m}{N_m^2}} = \frac{\beta}{2}. \tag{30}$$

Combining terms and rearranging yields:

$$\sum_{m \in I} \frac{\Gamma_m}{N_m^2} \left[\frac{(N_m - 1)^2}{N_m} - \frac{\beta}{2}\right] = 0. \tag{31}$$

Suppose the total market size Γ_m for each submarket are ranked from smallest to largest such that $\Gamma_0 \leq \Gamma_1 \leq \cdots \leq \Gamma_M$. Then by summation by parts, equation (31) can be written as:

$$\left[\frac{(N_M - 1)^2}{N_M} - \frac{\beta}{2}\right] \sum_{l=0}^{M} \frac{\Gamma_l}{N_l^2} - \sum_{l=0}^{M-1} \sum_{k=0}^{l} \frac{\Gamma_k}{N_k^2} \left[\frac{(N_{l+1} N_l - 1)(N_{l+1} - N_l)}{N_{l+1} N_l}\right] = 0. \tag{32}$$

Since $\sum_{l=0}^{M} \frac{\Gamma_l}{N_l^2}$ is equivalent to total firm profit, dividing equation (32) through by total profit obtains an expression in terms of the summation of the proportion ρ_m of total profit attributed to each submarket m such that: $\rho_m = \sum_{k=0}^{m} \frac{\Gamma_k}{N_k^2}$ Therefore, equation (32) can be written as:

$$\left[\frac{(N_M-1)^2}{N_M}-\frac{\beta}{2}\right]-\sum_{l=0}^{M-1}\rho_l\left[\frac{(N_{l+1}N_l-1)(N_{l+1}-N_l)}{N_{l+1}N_l}\right]=0. \tag{33}$$

If fixed costs are exogenous, then the monotonocity of equation (22) implies that the number of firms that enter in equilibrium in each submarket maintains the same ordering. Thus, $N_{m+1} \geq N_m, \forall m$ and the second term in equation (33) is non-negative such that $\left[\frac{(N_M-1)^2}{N_M}-\frac{\beta}{2}\right] \geq 0$. On the other hand, Sutton's EFC model (1991; 1998) implies that the presence of endogenous sunk costs limits the equilibrium number of firms that can enter even as the market size becomes large. Therefore, for some critical value of market size Γ_{m^*}, industries switch from being characterized by exogenous fixed costs to endogenous fixed costs such that $N_{m+1} \leq N_m, \forall m > m^*$. Provided that the number, and size, of submarkets characterized by endogenous fixed costs are greater than those characterized by exogenous fixed costs, the second term in equation (33) is non-positive such that $\left[\frac{(N_M-1)^2}{N_M}-\frac{\beta}{2}\right] \leq 0$. Thus, if the largest submarket M is characterized by endogenous (exogenous) fixed costs, then solving $\left[\frac{(N_M^*-1)^2}{N_M^*}-\frac{\beta}{2}\right]=0$ for N_M^* yields the least upper bound (greatest lower bound) to the number of firms that enter in equilibrium.

The number of firms entering under endogenous fixed costs N_m^{EN} solves $\left[\frac{(N_m-1)^2}{N_m}-\frac{\beta}{2}\right]=0$ which can be expressed equivalently as:

$$N_m^2-N_m\left(\frac{\beta}{2}+2\right)-1=0. \tag{34}$$

The roots to the quadratic equation (34) are equal to:

$$N_m^{EN} = \left(1 + \frac{\beta}{4}\right) \pm \frac{1}{4}\sqrt{\beta^2 + 4\beta}.$$

Given the assumption on the cost elasticity parameter $\beta > 2$, the smaller of the two roots is always less than one such that the equilibrium number of firms that enter under endogenous fixed costs equals:

$$N_m^{EN} = \frac{1}{4}(\beta + 4) + \frac{1}{4}\sqrt{\beta(\beta + 4)}. \qquad (35)$$

Figure 1 illustrates the lower bound to concentration under exogenous ($\underline{C}_{EX}$) and endogenous ($\underline{C}_{EN}$) fixed costs for sets of parameter values $\{F_0, \beta\}$. As the R&D cost parameter β increases, the upper limit on the total number of firms that enter in equilibrium increases such that the lower bound to concentration $\underline{C}_{EN}$ under endogenous fixed costs decreases. Moreover, as the minimum setup cost F_0 increases, the total number of firms that enter in equilibrium decreases, hence shifting the lower bound to concentration $\underline{C}_{EX}$ under exogenous fixed costs outward.

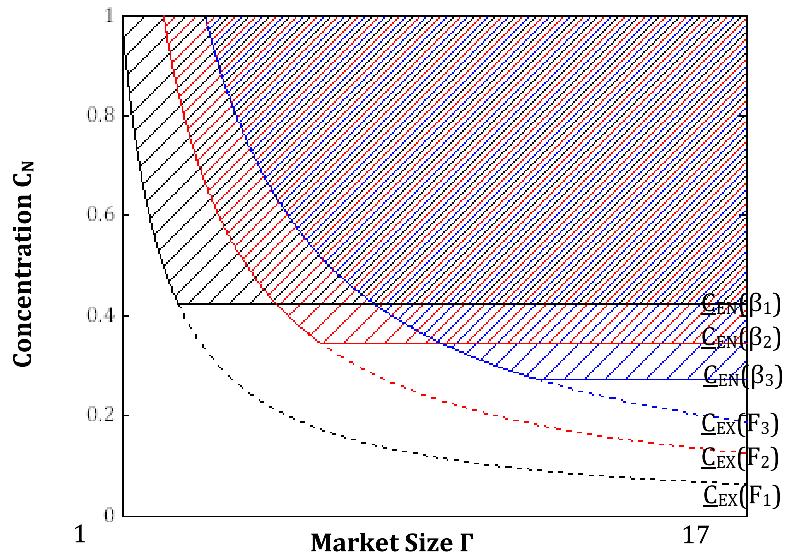


Figure 1: Illustrating equilibrium concentration and market size from the Cobb-Douglas demand example (under symmetric equilibrium and simultaneous entry). The R&D cost parameter $\beta_3>\beta_2>\beta_1$ illustrates the greater lower bound under lower R&D costs relative to the lower bound under higher R&D costs. The minimum setup cost $F_3>F_2>F_1$ illustrates the horizontal shift of firm entry associated with greater exogenous sunk costs.

Figure 1: Equilibrium Concentration Levels and Market Size

By equating the equilibrium number of firms from equations (22) and (35), we determine the threshold value for the market size Γ_{m^*} whereby a submarket m changes from being characterized by exogenous fixed costs to endogenous fixed costs. Specifically,

$$\Gamma_{m^*} = F_0\left[\frac{1}{4}(\beta+4)+\frac{1}{4}\sqrt{\beta^2+4\beta}\right]^2. \tag{36}$$

The upshot from this analysis is that for sufficiently sized markets, the ability of firms to increase quality via fixed (sunk) R&D investments precludes additional entry by new firms such that existing firms capture further expansions of market size via quality escalation. Thus, even as the size of the market grows large (i.e. $\Gamma_{m^*}\to\infty$) firm concentration levels remain bounded away from perfectly competitive levels (i.e. $C_1 \gg 0$).

A Lower Bound to R&D Concentration

The illustrative model developed in the previous section relates the number of firms that enter in equilibrium, hence industry concentration, to total market size and the endogeneity or exogeneity of sunk R&D expenditures. Sutton (1998) finds that the lower bound to R&D intensity, measured as the ratio of firm R&D to firm sales, is equivalent to the lower bound to concentration as markets become large. However, he does not address the implications of the EFC model upon concentration of R&D within these industries, which remains as a separate and additional concern in discussions regarding mergers and acquisitions, as well as patent pools, in agricultural biotechnology (Moschini, 2010; Dillon and Hubbard, 2010; Moss, 2009). 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.

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 submarket m can be stated as:

$$C_{1m} = \frac{\widehat{\Pi}_m}{\Pi_m} \geq \alpha(\sigma, \beta) \cdot h_m, \tag{37}$$

where α is some constant for a given set of parameter values (σ, β) and is independent of the size of the market in endogenous fixed cost industries. The value of alpha α depends upon industry technology, price competition, and consumer preferences and captures the extent that a firm can escalate quality via R&D investment and capture greater market share from rivals.

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 submarket m, we define total industry sales revenue Π_m and R&D expenditure F_m in submarket 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 submarket m accounted for by the largest product category such that:

$$h_m = \max_l \frac{\Pi_{lm}}{\Pi_m}, \tag{38}$$

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

Moreover, from Theorem 3.2 implies an equivalent expression for the lower bound to R&D-intensity $\widehat{\mathrm{P}}_m$ for the quality-leading firm such that:

$$\hat{\mathrm{P}}_m = \frac{\hat{F}_m}{\hat{\Pi}_m} \geq \alpha(\sigma, \beta) \cdot h_m - \frac{F_0}{\Pi_m}. \tag{39}$$

Equation (39) 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. $\Pi_m \to \infty$). Multiplying both sides of equation (38) by $\hat{\Pi}_m$ yields:

$$\hat{F}_m \geq \alpha(\sigma, \beta) \cdot h_m \cdot \hat{\Pi}_m - F_0 \cdot \frac{\hat{\Pi}_m}{\Pi_m}. \tag{40}$$

Dividing both sides of equation (40) by total industry sales revenue in submarket m yields:

$$\frac{\hat{F}_m}{\Pi_m} \geq \left[\alpha(\sigma, \beta) \cdot h_m - \frac{F_0}{\Pi_m}\right] \cdot \frac{\hat{\Pi}_m}{\Pi_m}. \tag{41}$$

However, free entry in equilibrium implies that total industry sales revenue Π_m equals total industry R&D expenditure F_m such that equation (41) can be written as:

$$\frac{\hat{F}_m}{F_m} \geq \left[\alpha(\sigma, \beta) \cdot h_m - \frac{F_0}{\Pi_m}\right] \cdot \frac{\hat{\Pi}_m}{\Pi_m}. \tag{42}$$

Defining the ratio of R&D concentration for the quality-leading firm as $R_{1m} = \frac{\hat{F}_m}{F_m}$, substituting for condition (37) on the lower bound to the single-firm concentration ratio, and substituting observable market size Γ_m for profit Π_m yields:

$$R_{1m} \geq \left[\alpha^2(\sigma, \beta) h_m^2 - \alpha(\sigma, \beta) h_m \frac{F_0}{\Gamma_m}\right]. \tag{43}$$

Equation (43) 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 (43) 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 submarket m can be expressed as:

$$R_{1m} = \frac{\hat{F}_m}{F_m} = \frac{F_0}{\Gamma_m}. \tag{44}$$

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 illustrates the relationship between R&D concentration and market size for both endogenous and exogenous fixed cost industries.

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 Γ increases. If an industry is characterized by homogenous products (i.e. low h), 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) 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.

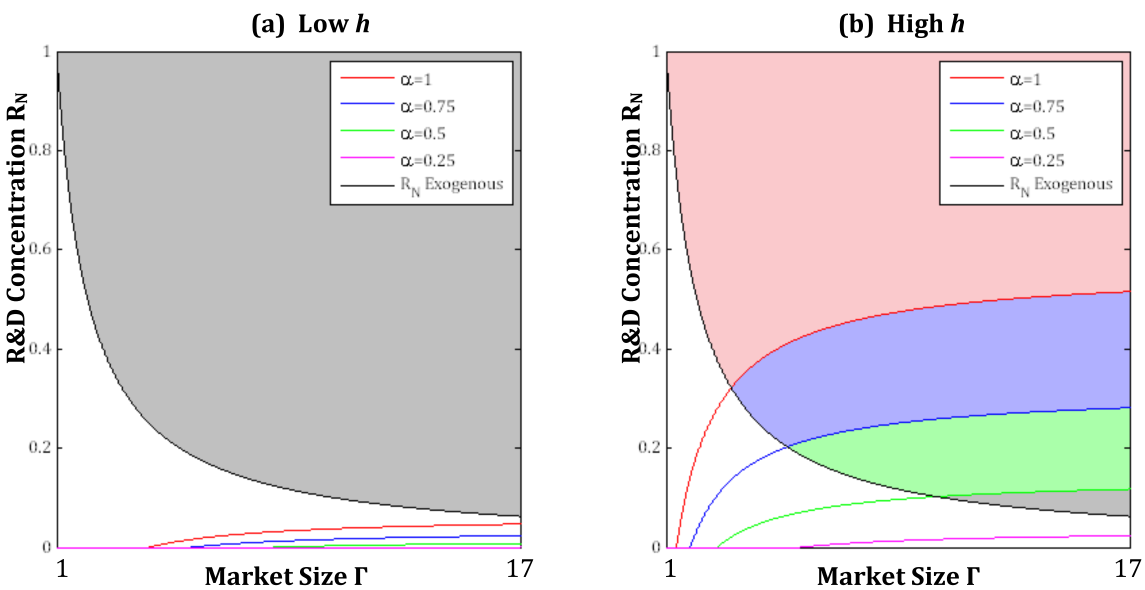


Figure 2: Equilibrium R&D Concentration Levels and Market Size

Figure 2: Illustrating equilibrium R&D concentration levels and market size for industries with low (2.a) and high (2.b) levels of product heterogeneity. For homogenous product industries (low *h*), no value of α would permit endogenous fixed costs such that R&D concentration decreases with market size. For heterogeneous product industries (high *h*), provided α is sufficiently large, R&D concentration will remain bounded away from zero as market size increases.

Empirical Specification

Equations (43) and (44) 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:[3]

$$\tilde{R}_n = \ln \frac{1}{1 - R_n}. \tag{45}$$

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

$$\frac{\tilde{R}_{nm}}{h_m^2} = \theta_0 + \theta_1 \frac{1}{h_m \ln(\Gamma_m / F_0)} + \varepsilon_m, \tag{46}$$

where the residuals ε between the observed values of R&D concentration and the lower bound are distributed according the Weibull distribution such that:

$$F(\varepsilon) = 1 - \exp\left[-\left(\frac{\varepsilon - \mu}{\delta}\right)^{\gamma}\right], \quad \gamma > 0, \delta > 0 \tag{47}$$

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

To test for a lower bound to R&D concentration, it is equivalent to test whether the residuals fit a two or three parameter Weibull distribution, that is to test whether $\mu = 0$. However, as Smith (1985) identifies, fitting equation (46) directly via maximum likelihood

[3] 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.

estimation is problematic for shape parameter values $\gamma \leq 2$.[4] 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:

$$\min_{\{\theta_0,\theta_1\}} \sum_{m=1}^{N} \left[\frac{\tilde{R}_{nm}}{h_m^2} - \left(\theta_0 + \theta_1 \frac{1}{h_m \ln(\Gamma_m/F_0)}\right)\right]$$

$$s.t. \qquad \frac{\tilde{R}_{nm}}{h_m^2} \geq \left(\theta_0 + \theta_1 \frac{1}{h_m \ln(\Gamma_m/F_0)}\right), \forall m.$$

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:

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

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 cannot be rejected, then this implies the presence of a

[4] Specifically, for $1 < \gamma \leq 2$, the 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.

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 a 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 for each (sub-) market. 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.

Sutton (1998) identifies the potential use of “natural experiments” in order to empirically identify the lower bound to concentration within a single industry. The natural experiments that allow for such an analysis occur when there is an exogenous shift in consumer preferences or an exogenous change in technology, although exogenous changes in market size also prove useful for analysis. 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 submarkets as well as over time. In doing so, we are able to capitalize upon changes in 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 intertemporal variation in market size permits the theory to be tested across a variety of market sizes. Finally, we are able to utilize a “natural experiment”, in the form of differential changes in demand for corn (positive) and soybean (negative) seed in response to an exogenous increase in the incentives for farmers to grow corn crops for use in ethanol.

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 is unavailable for the agricultural biotechnology sector, there is publicly available data that captures 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 is 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 data are appropriate for the analysis as it captures 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). These Field Trial Applications 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 9936 remaining observations in the database.

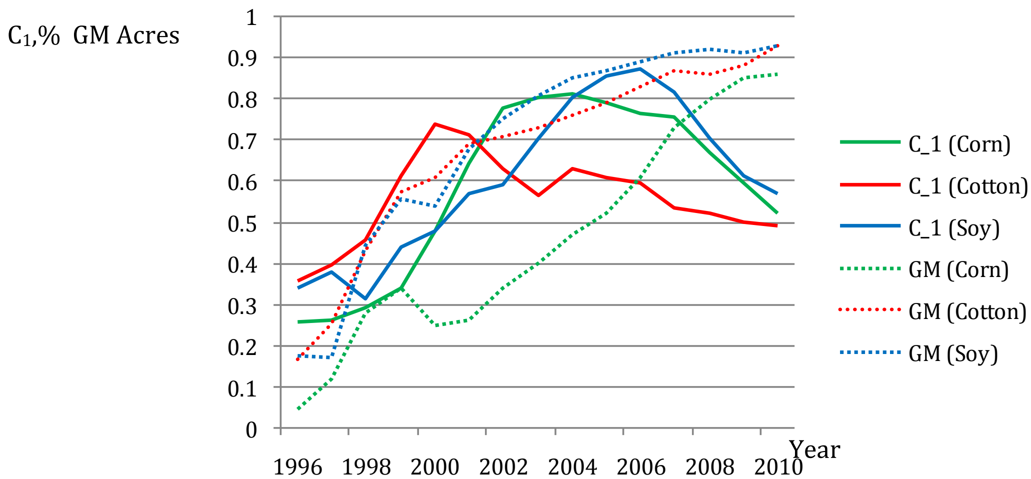
The National Agricultural Statistics Service (NASS), a division of the United States Department of Agriculture (USDA), conducts the annual June Agriculture Survey in order to obtain estimates of farm acreage for a variety of crops, including corn, cotton, and soybeans. The NASS reports data on total 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.

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 this survey data, ERS computes and reports estimates for the extent of GM adoption. GM adoption rates 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. Figure 3 plots 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 concentration ratios that initially increased and have remained consistently high across time.

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

[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. Provided rates of adoption or total planted acreage are not significantly greater among these "marginal" states, this imputation will not bias the estimates.

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. Therefore, we treat seed varieties as homogenous within product groups, 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.



Source: Author's calculations from APHIS and ERS GM crop adoption data.

Figure 3: Single-Firm R&D Concentration Ratios and GM Adoption

The final component required for the estimation of the lower bound to R&D concentration is the minimum setup cost associated with entry into the product market. We use data reported in the "National Plant Breeding Study" for 1994 and 2001 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 divide this sum by the total number of projects reported for both agencies in order to obtain average SY for a single crop.[6] Minimum setup costs are thus obtained by multiplying average SY by the private industry cost per SY ($148,000) and adjusting for inflation.[7] Table 3 reports summary statistics for field trial applications, crop acreage planted, seed costs, product heterogeneity, and minimum setup costs.

Table 1: Lower Bound Estimation Data Descriptive Statistics

Variable	Yearly Mean	Std. Dev.	Min.	Max.	Total
Field Trial Applications					
Corn	318.86	166.46	3.00	606.00	6696.00
Cotton	41.13	23.63	1.00	91.00	946.00
Soybeans	78.32	55.94	4.00	194.00	1723.00
Total Acres Planted (000 acre)					
Corn	79945.43	5176.88	71245.00	93600.00	1678854.00
Cotton	13434.61	1994.88	9149.50	16931.40	282126.90
Soybeans	69462.76	7287.88	57795.00	82018.00	1458718.00
Seed Costs ($/acre)					
Corn	25.13	7.10	15.65	49.15	-
Cotton	25.46	15.63	7.24	79.55	-
Soybeans	18.34	6.99	7.79	37.28	-
Product Heterogeneity					
Corn	0.69	0.24	0.35	1.00	-
Cotton	0.59	0.26	0.31	1.00	-
Soybeans	0.84	0.14	0.64	1.00	-
Minimum Setup Costs ($1000)					
Corn	182536.55	27275.81	147843.26	220132.27	-
Cotton	182019.44	27434.03	147424.44	219508.67	-
Soybeans	282336.98	42553.92	228675.42	340487.88	-

Source: Author's estimates

[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] Results of the lower bound estimations are robust to an alternate definition based upon public sector cost per SY ($296,750).

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 4) 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. 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.

Table 2: Market Definition Data Descriptions

Observable Market Characteristics			
State Level			
Data	**Description**	**Years**	**Source**
Latitude	State geographic centroid	-	MaxMind®
Longitude	State geographic centroid	-	MaxMind®
Size	Total area (000s acres)	-	2000 Census of Population and Housing
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

V. 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 (2003), 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 first 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. Moreover, recent surveys of global agricultural biotechnology indicate that many of the varieties of GM crops adopted outside of the US have also been developed outside of the US. (ISAAA, 2010)

We consider a characterization of regional submarkets 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 submarkets as it captures the "natural structure" of the data across multiple characteristics. 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.

The goal of cluster analysis is such that objects within a cluster (i.e. states within a regional submarket) are “close” in terms of observable characteristics while being “far” from objects in other clusters. Thus, the objective is to define distinct, exclusive submarkets in the agricultural seed sector by clustering states into non-overlapping partitions. We assume a “prototype-based” framework such that every state in some submarket is more similar to some prototype state that characterizes its own submarket relative to the prototype states that characterize other submarkets. Therefore, we utilize a K-means approach by defining the number of submarket 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.

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. submarkets). 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 Stiegert, Shi, and Chavas (2011) by separating the major corn production regions into "core" and "fringe" states and refining 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 pesticide applications prior to the introduction of GM crops. The resulting submarkets, summarized in Figure 3 with the submarket shares of total US production, reveals that corn production is heavily concentrated in only thirteen states with Illinois and Iowa alone accounting for approximately 30% of all production.

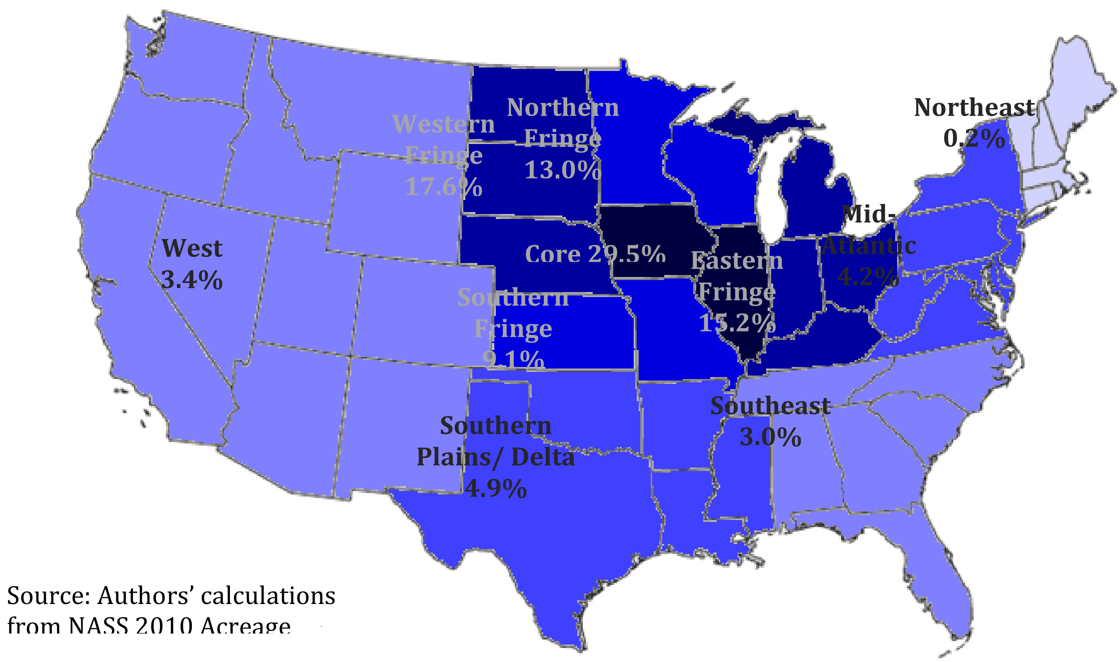


Figure 4: 2010 Submarket Shares of US Corn Acres Planted

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 submarkets apiece. However, there are large differences in the relative size of submarkets 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 (Figure 5). Soybean production, on the other hand, primarily occurs in corn-producing regions with the significant overlap between the major corn and soybean producers (Figure 6).

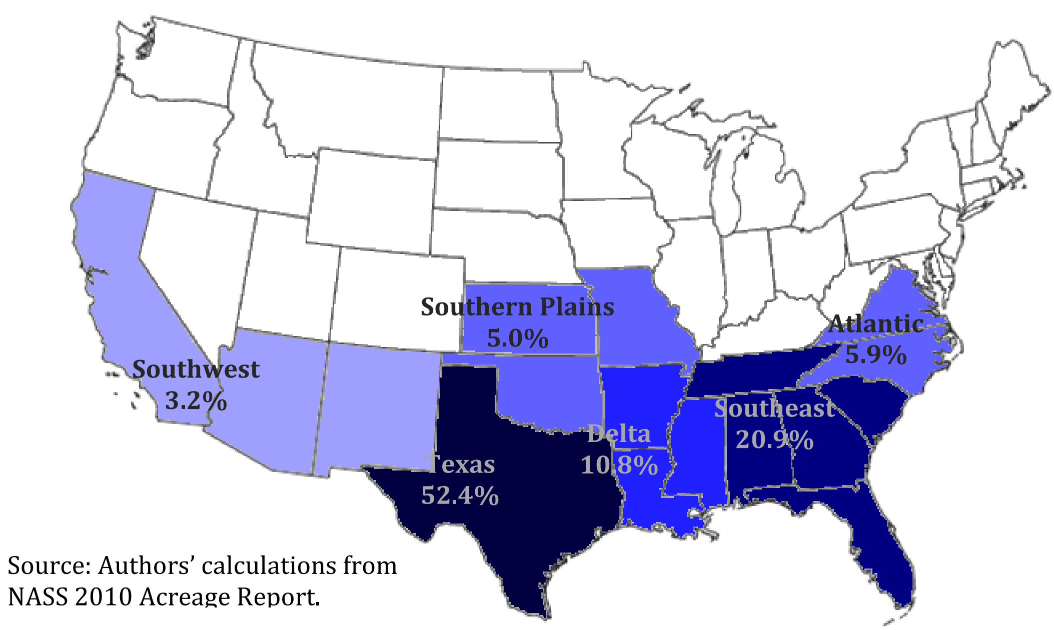
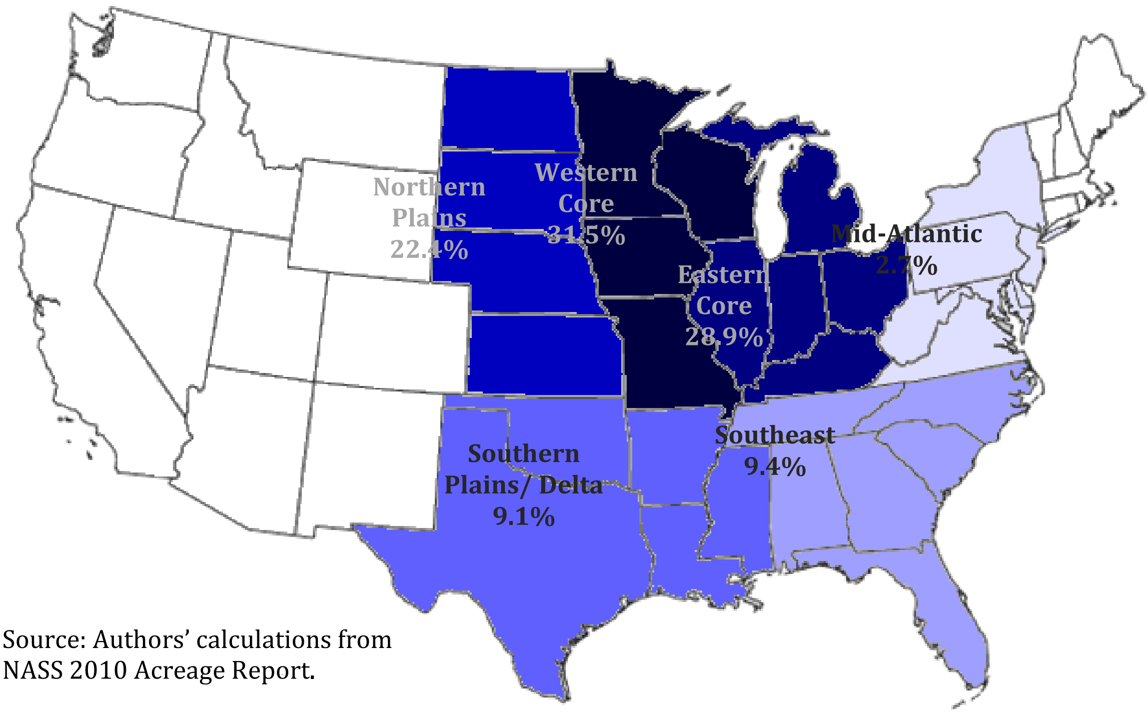


Figure 5: 2010 Submarket Shares of US Cotton Acres Planted



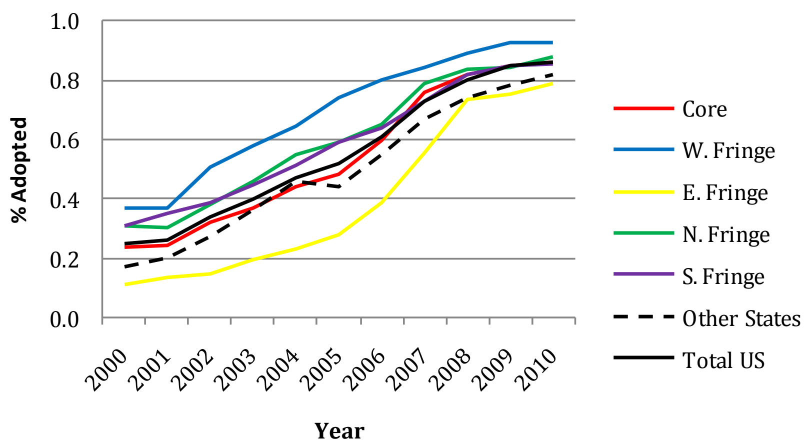
Source: Authors' calculations from NASS 2010 Acreage Report.

Figure 6: 2010 Submarket Shares of US Soybean Acres Planted

Examining rates of GM adoption across submarkets (Figures 7-9) reveals distinct differences across submarkets for GM corn and cotton seeds, but similar adoption rates for GM soybean. One possible explanation for the observed differences across crop types might be the relatively limited number of phenotypes of GM soybean released in this period such that all regions had a similar preference for insect resistance.[8] Although there is only limited number of states with available data on the percentage of farm acres planted with soybeans and treated with insecticide from 1990-1995, only two states (Georgia and

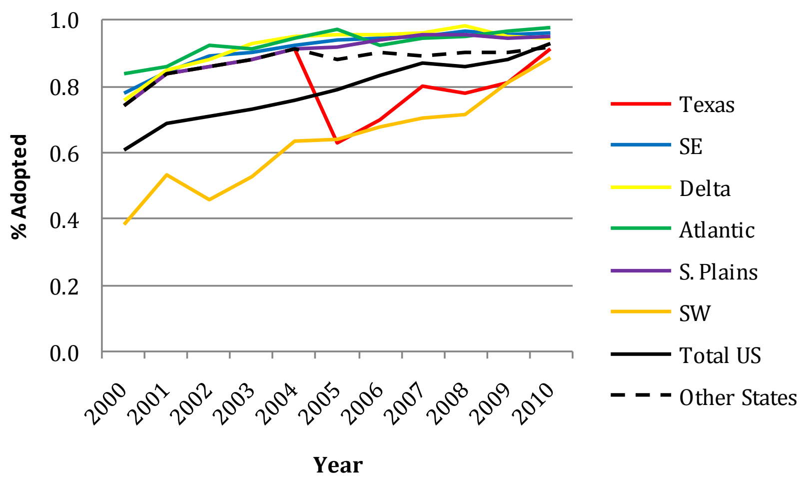
[8] For additional descriptive analysis of the geographically distinct submarkets, please refer to "Appendix A: (Sub-)Market Analysis for GM Crops". Specifically, Appendix A contains maps illustrating the geographical differences in climate and market size for each crop type as well as differences in the application of fertilizers, herbicides, and insecticides by crop.

Louisiana) reported percentage of acres treated at greater than 10% with the remaining states (Arkansas, Illinois, Indiana, Kentucky Mississippi, Missouri, Nebraska, North Carolina, Ohio, South Dakota) reporting rates that were typically less than 5% of acres treated. These descriptive statistics contrast greatly with those for cotton, which also had a limited number of phenotypes released in this period. Texas, which had the lowest rates of adoption of Bt Cotton, an insect resistant variety, also had the lowest rates of insecticide application from 1990-1995.



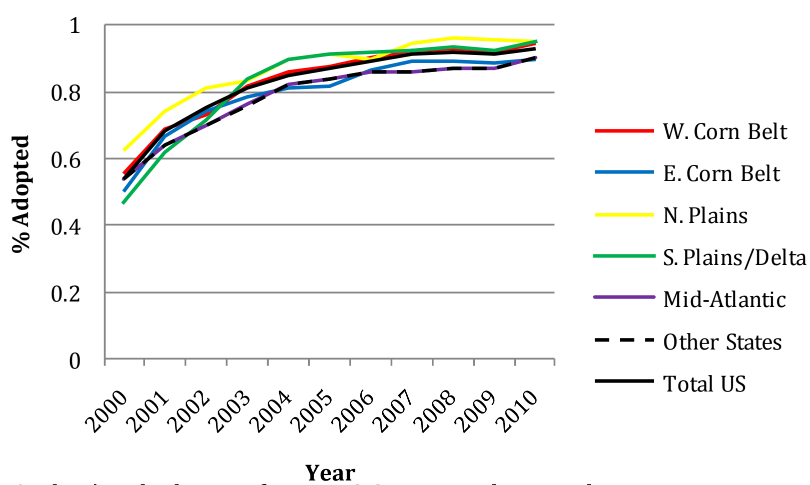
Source: Author's calculations from ERS GM crop adoption data.

Figure 7: Adoption Rates of GM Corn Across Submarkets



Source: Author's calculations from ERS GM crop adoption data.

Figure 8: Adoption Rates of GM Cotton Across Submarkets



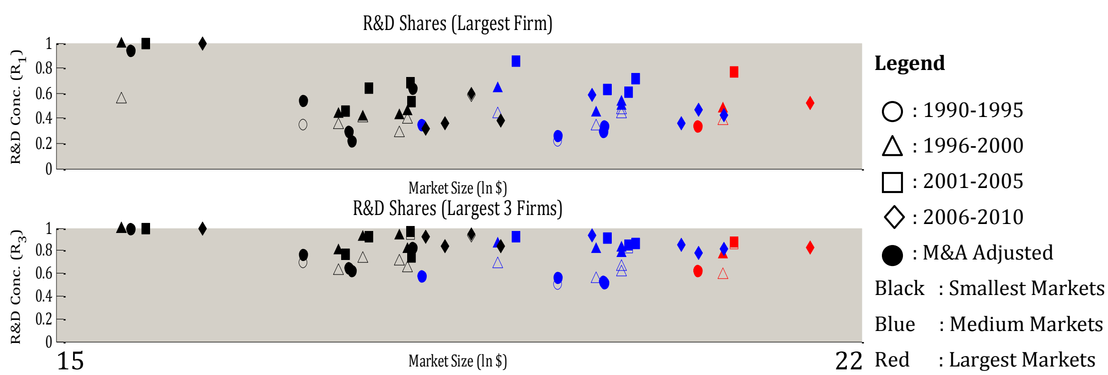
Source: Author's calculations from ERS GM crop adoption data.

Figure 9: Adoption Rates of GM Soybean Across Submarkets

VI. Empirical Results and Discussion

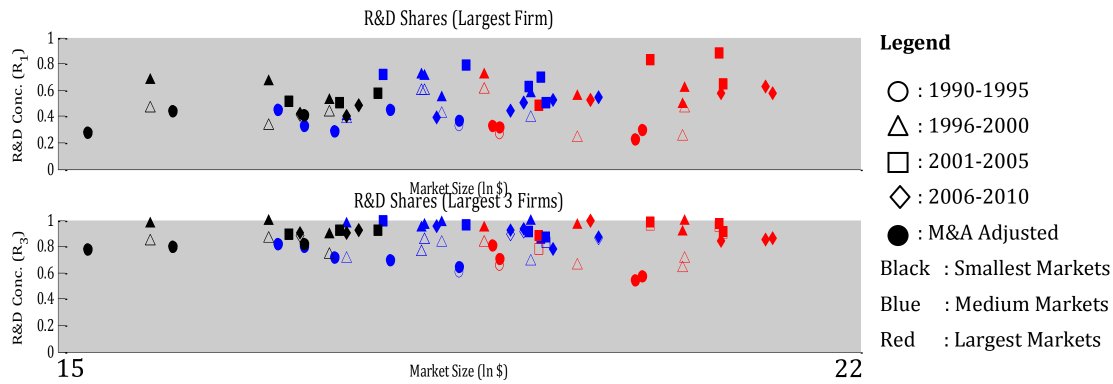
Estimating the Lower Bounds to R&D Concentration

Prior to estimating the lower bounds to R&D concentration, it is useful to illustrate why one would expect the agricultural biotechnology sector to be characterized by endogenous lower bounds. Specifically, Figures 7-9 illustrate the one- and three-firm R&D concentration ratios relative to market size for each crop type with and without adjustments for merger and acquisition activity.



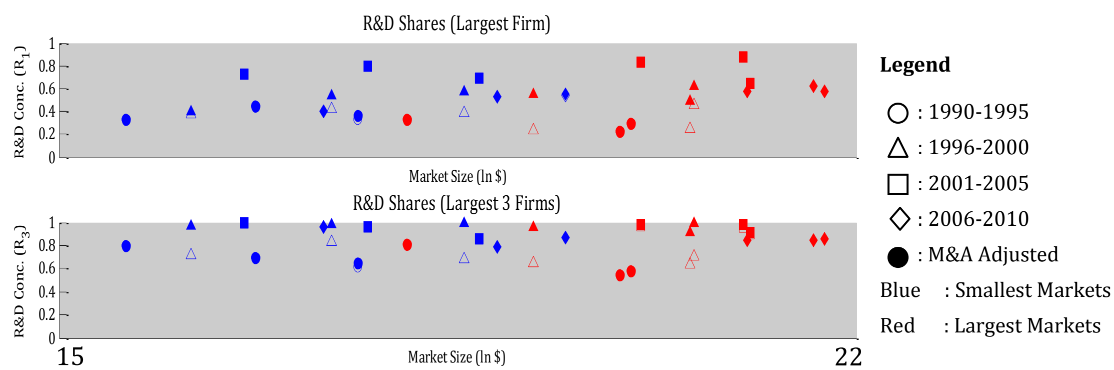
Source: Authors' calculations.

Figure 10: Corn R&D Concentration and Market Size



Source: Authors' calculations.

Figure 11: Cotton R&D Concentration and Market Size



Source: Authors' calculations.

Figure 12: Soybean R&D Concentration and Market Size

Figures 10-12 illustrate that R&D concentration ratios are non-decreasing in market size for each type of crop regardless of the measure of R&D concentration, therefore implying 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 with the lower bound to R&D concentration implied by the theory. As such, the following analysis considers eight variations for each crop type by analyzing both the single and three-firm concentration, concentration adjusted and unadjusted for merger and acquisition of intellectual property, and both total market size for each crop as well as the market size for genetically engineered crops (1996-2000, 2001-2005, 2006-2010). The two-stage estimation results as well as illustrative figures are presented for each crop type.

R&D Concentration in GM Corn Seed

The lower bounds to concentration for corn seed are illustrated in Figure 13 and the estimation results are presented in Table 5. The results indicate a lower bound to R&D concentration that is increasing in the size of the market, independent of the definitions of R&D concentration and market size. 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, which contrasts with the exogenous lower bound to R&D concentration which is strictly decreasing in market size. Moreover, factoring merger and acquisition activity into the measurement of R&D concentration does not significantly change the estimates for the lower bound to concentration in corn seed. These results imply that increased concentration of intellectual property in corn seed occur not as a consequence of merger and acquisition activity, but rather are inherent in the nature of technological competition.

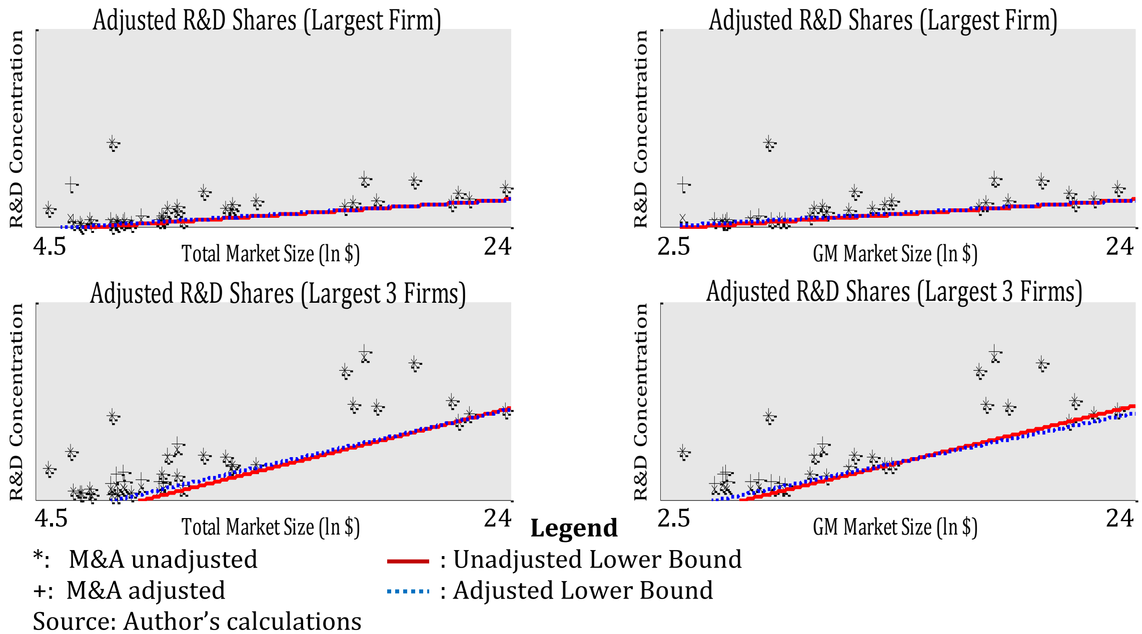


Figure 13: Lower Bounds to R&D Concentration in GM Corn Seed

When interpreting the results of the lower bound estimation, it is important to recall that the R&D concentration data have been adjusted according to equation (45) and for the level of product heterogeneity. For large markets, the predicted lower bound converges to 1 whereas for very small markets, the predicted lower bound converges to some negative number (i.e. 0). This result can be explained by what we observe in Figures 2 and 10. According to the theory, we are attempting to fit a lower bound to an industry that is characterized by exogenous fixed costs in smaller-sized markets and endogenous fixed costs in larger-sized markets such that there is a structural break for some unidentified level of market size. A more accurate estimation based upon the observed data could fit a non-linear lower bound to the data to account for this structural break, although it is difficult *a priori* to reconcile such an analysis with the theoretical predictions. Moreover, Figure 13, which plots the adjusted data, does not necessarily indicate that a non-linear lower bound would provide a better fit.

In order to interpret the coefficient estimates over the range of possible market sizes, we consider 10% changes in the market size for both the largest and smallest submarkets and report the predicted lower bound results in Table 8. In the largest submarket (Iowa and Illinois), the predicted lower bound of the single firm R&D concentration ratio ranges from .3776 to .3909 and a 10% increase in market size increases the lower bound in the range of .0035 to .0049. For the smallest submarket (Northeast states), the range of predicted single firm R&D concentration ratios range from .1122 to .2178 and a 10% increase in market size increases the lower bound by a range from about .0044-.0071. The predicted values of the three firm R&D concentration ratios range from .7887 to .8152 in large markets and .2409 to .5062 for small markets. A 10% increase in market size increases the lower bound in small markets between .0102-.0226 and increases the lower bound in large markets between .0044-.0055.

Table 3: Lower Bound Estimations for GM Corn Seed

Concentration	Market Size	M&A	First-Stage θ_0	First-Stage θ_1	Second-Stage γ	Second-Stage δ	LR (χ^2=1)
R_1	Total	Unadjusted	-1.714 **	0.240 **	0.747 **	1.925 **	0.161
			0.060	0.006	0.080	0.445	
		Adjusted	-1.444 **	0.228 **	0.718 **	1.913 **	0.117
			0.042	0.005	0.079	0.459	
	GM	Unadjusted	-0.443 **	0.186 **	0.710 **	2.126 **	-0.006
			0.034	0.004	0.091	0.601	
		Adjusted	-0.069 *	0.168 **	0.690 **	2.118 **	0.034
			0.030	0.004	0.090	0.617	
R_3	Total	Unadjusted	-7.804 **	0.901 **	1.007 **	4.975 **	-0.020
			0.148	0.011	0.119	0.846	
		Adjusted	-6.244 **	0.830 **	0.946 **	4.403 **	0.016
			0.119	0.010	0.111	0.801	
	GM	Unadjusted	-3.211 **	0.725 **	0.770 **	4.007 **	0.022
			0.083	0.007	0.112	1.038	
		Adjusted	-1.923 **	0.625 **	0.813 **	4.273 **	-0.030
			0.139	0.012	0.116	1.026	

Source: Author's estimates.

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

R&D Concentration in GM Cotton Seed

As with the market for corn seed, the estimation of a lower bound to R&D concentration in cotton seed implies an industry characterized by endogenous fixed costs to R&D. Figure 14 illustrates lower bounds to R&D concentration that are again increasing in market size in each estimation. The results, reported in Table 6, imply a significant and increasing lower bound to R&D concentration that is not independent of the size of the market (i.e. $\theta_1 > 0$). However, when merger and acquisition activity are accounted for in R&D concentration, the predicted lower bound for the cotton seed market changes significantly in the three-firm concentration estimations, thus implying some of the observed concentration in intellectual property in cotton seed has occurred as a result of firm mergers and acquisitions and cannot necessarily be attributed to the nature of technology competition.

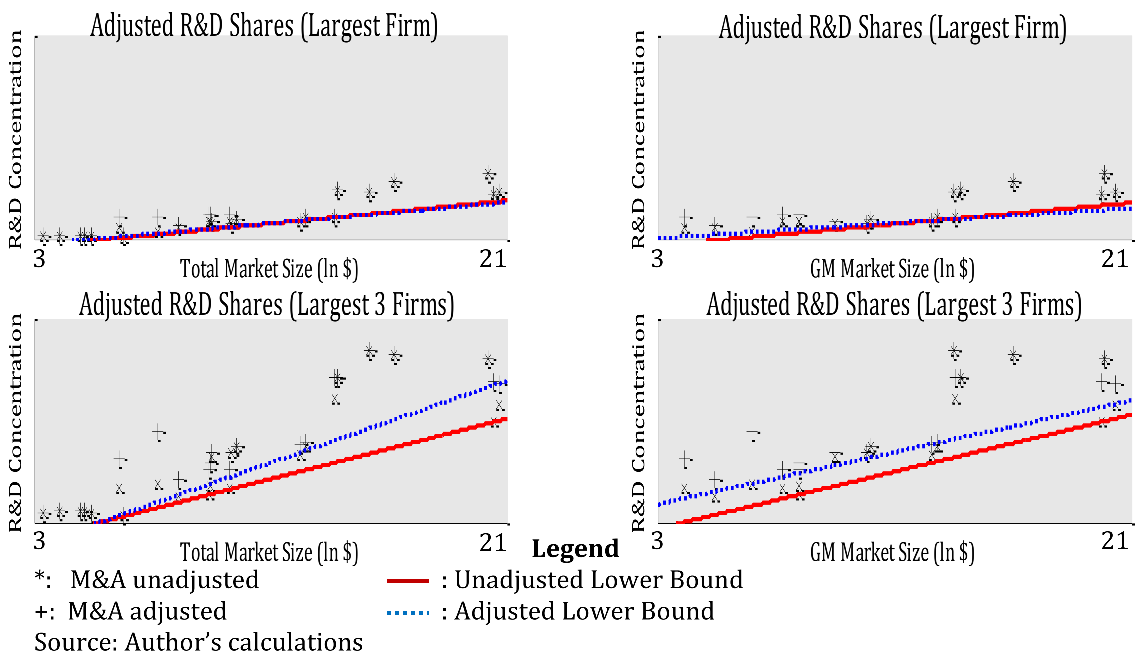


Figure 14: Lower Bounds to R&D Concentration in GM Cotton Seed

The predicted lower bound to single firm R&D concentration for GM cotton seed range from .3539 to .4302 for the largest sized market (Texas) and from .2388 to .2662 for the smallest sized market (Southwest states). A 10% increase in market size increases the lower bound to concentration from between .0045 to .0081 regardless of the size of the market. Furthermore, the predicted lower bound for the three firm R&D concentration ratios range from .7811 to .8767 in the largest market with a 10% increase in market size raising the predicted lower bound by .0042-.0063. In the smallest-sized market, the predicted lower bound to the three firm R&D concentration ratio ranges from .5765 to .6917 and a 10% increase in market size increases the predicted lower bound by .0074 to .0123.

Table 4: Lower Bound Estimations for GM Cotton Seed

			First-Stage				Second-Stage				LR
Concentration	Market Size	M&A	θ_0		θ_1		γ		δ		$(\chi^2=1)$
R_1	Total	Unadjusted	-2.203	**	0.372	**	1.260	**	1.579	**	0.028
			0.121		0.010		0.195		0.277		
		Adjusted	-1.822	**	0.343	**	1.405	**	1.947	**	0.066
			0.123		0.010		0.231		0.313		
	GM	Unadjusted	-1.606	**	0.328	**	1.492	**	2.277	**	0.058
			0.287		0.022		0.281		0.391		
		Adjusted	-0.366		0.233	**	1.474	**	2.551	**	0.010
			0.335		0.026		0.289		0.442		
R_3	Total	Unadjusted	-5.459	**	0.986	**	1.114	**	5.166	**	-0.008
			0.432		0.042		0.188		1.040		
		Adjusted	-7.426	**	1.347	**	1.400	**	5.101	**	0.008
			0.268		0.019		0.234		0.821		
	GM	Unadjusted	-2.715	**	0.881	**	1.328	**	6.181	**	0.114
			0.428		0.040		0.255		1.234		
		Adjusted	1.037	*	0.803	**	1.022	**	4.656	**	-0.018
			0.507		0.041		0.203		1.160		

Source: Author's estimates.

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

Comparing the predicted lower bounds of the corn and cotton seed markets, it is evident that the lower bound to R&D concentration in cotton seed increases somewhat more rapidly relative to corn seed. This result can be explained in part by the proliferation of products in the GM corn seed market (29 GM seed varieties) relative to the number of GM cotton seeds marketed (11 GM seed varieties). (Howell, et al., 2009) As the h-index decreases with the level of product heterogeneity, the R&D concentration for GM cotton seed is more likely to be characterized by endogenous fixed costs.

R&D Concentration in GM Soybean Seed

Unlike the estimations for GM corn and GM cotton seed, the lower bound estimations for R&D concentration in GM soybean seed, reported in Figure 15 and Table 10, are more ambiguous. Although six of eight lower bound estimations are indicative of endogenous fixed costs in R&D with R&D concentration increasing with market size, the estimations for the GM market are relatively flat and it is not evident that these estimations are significantly different from zero for feasible market sizes. Moreover, from the estimations for total market size, the results indicate that even though the market appears to be characterized by endogenous fixed costs, much of the concentration has occurred as a result of merger and acquisition activity. Despite relatively high levels of product homogeneity (5 GM soybean varieties) and data points that imply a lower bound to R&D concentration in Figure 12, the empirical results provide only partial evidence that the market for GM soybeans is characterized by endogenous fixed costs. (Howell, et al., 2009)

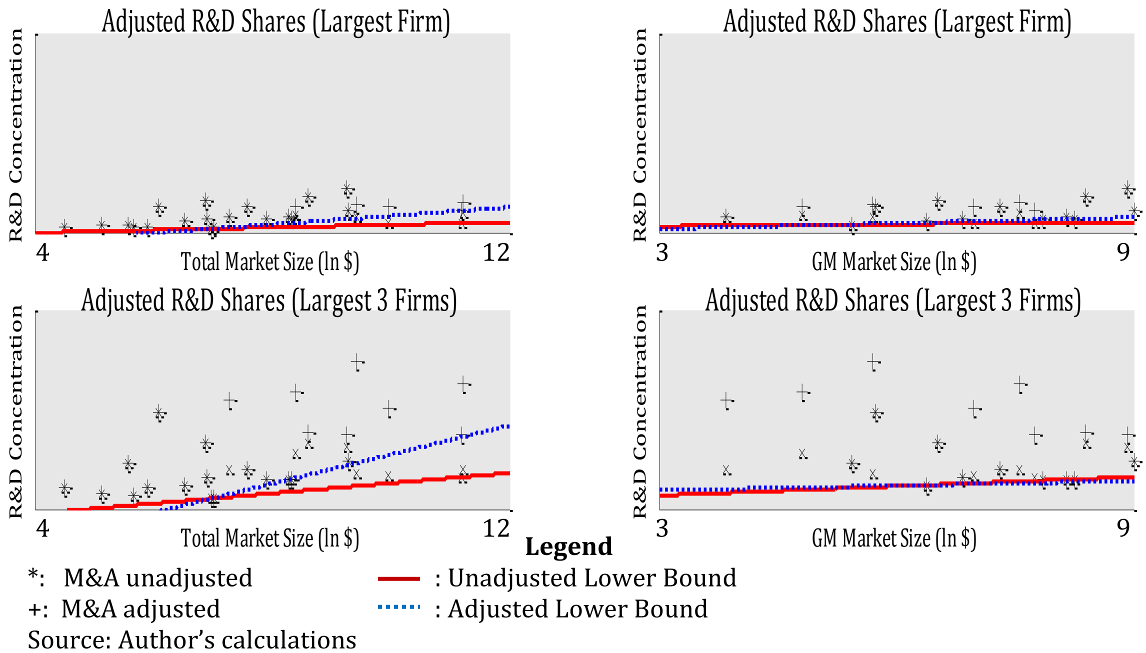


Figure 15: Lower Bounds to R&D Concentration in GM Soybean Seed

The predicted lower bound to R&D concentration for GM soybean seeds reveals the increase in R&D concentration resulting from mergers and acquisitions more clearly than either the estimations for GM corn or GM cotton seed. Whereas the lower bound to single firm R&D concentration ranges from .2625-.2998 in the unadjusted estimations for the largest-sized market, after adjusting for merger and acquisition activity the predicted lower bound increases to .4616-.5264. Moreover, a 10% increase in market size increases the predicted lower bound twice as rapidly when consolidation of intellectual property via mergers and acquisitions is considered. The estimations for the three firm R&D concentration ratios imply predicted lower bounds that range from .3539 for the smallest-sized market to .9187 for the largest-sized market. Contrary to the results for the single firm R&D concentration ratio, the predicted lower bound when considering mergers and acquisitions is lower.

It is also important to address the similarities and differences in the second-stage estimations for GM corn, cotton, and soybean seeds. 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. Additionally, the scale parameter δ describes the dispersion of the data. Most interesting are the results on the shape parameter γ which imply a high degree of clustering on the lower bound for all crop types, with cotton being characterized by the least clustering and corn the most. Moreover, γ is less than two in all 24 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 the three-firm estimations relative to the one-firm estimations.

Table 5: Lower Bound Estimations for GM Soybean Seed

			First-Stage				Second-Stage				LR
Concentration	Market Size	M&A	θ_0		θ_1		γ		δ		(χ^2=1)
R_1	Total	Unadjusted	-0.450	**	0.100	**	1.325	**	1.015	**	0.006
			0.105		0.014		0.217		0.172		
		Adjusted	-1.990	**	0.321	**	1.674	**	1.092	**	-0.035
			0.137		0.017		0.272		0.147		
	GM	Unadjusted	0.302		0.048		1.261	**	0.963	**	-0.029
			0.191		0.027		0.249		0.201		
		Adjusted	-0.197		0.145	**	1.410	**	1.067	**	-0.060
			0.127		0.017		0.303		0.198		
R_3	Total	Unadjusted	-1.892	**	0.383	**	1.317	**	2.093	**	-0.033
			0.188		0.024		0.213		0.357		
		Adjusted	-6.835	**	1.091	**	0.913	**	2.888	**	-0.053
			0.332		0.041		0.168		0.699		
	GM	Unadjusted	0.317		0.232	**	1.009	**	1.609	**	-0.057
			0.371		0.049		0.205		0.418		
		Adjusted	1.048		0.117		1.082	**	3.727	**	-0.029
			0.863		0.119		0.221		0.876		

Source: Author's estimates.

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

Table 6: Predicted Lower Bounds for GM Corn, Cotton, and Soybean Seeds

			1 Firm R&D Concentration				3 Firm R&D Concentration			
			Total Market		GM Market		Total Market		GM Market	
			Merger & Acquisition		Merger & Acquisition		Merger & Acquisition		Merger & Acquisition	
			Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Corn	Largest Market	Bound	0.3909	0.3888	0.3807	0.3776	0.8152	0.8125	0.8141	0.7887
		10% Change	0.0049	0.0046	0.0039	0.0035	0.0055	0.0051	0.0045	0.0044
	Smallest Market	Bound	0.1122	0.1264	0.2025	0.2178	0.2409	0.3103	0.5023	0.5062
		10% Change	0.0071	0.0066	0.0050	0.0044	0.0226	0.0189	0.0120	0.0102
Cotton	Largest Market	Bound	0.4302	0.4166	0.3983	0.3539	0.7835	0.8767	0.7811	0.8228
		10% Change	0.0063	0.0060	0.0059	0.0045	0.0063	0.0049	0.0057	0.0042
	Smallest Market	Bound	0.2662	0.2635	0.2419	0.2388	0.5765	0.6917	0.5933	0.6884
		10% Change	0.0081	0.0075	0.0074	0.0053	0.0123	0.0122	0.0106	0.0074
Soybean	Largest Market	Bound	0.2625	0.5264	0.2998	0.4616	0.6631	0.9187	0.7140	0.6323
		10% Change	0.0045	0.0092	0.0021	0.0048	0.0078	0.0053	0.0040	0.0026
	Smallest Market	Bound	0.1251	0.1818	0.2422	0.3162	0.3539	0.4790	0.5808	0.5539
		10% Change	0.0128	0.0378	0.0022	0.0060	0.0354	0.0773	0.0059	0.0032

Source: Author's estimates

VII. Conclusions

In the second essay, we examine whether a specific industry, agricultural biotechnology, is characterized by endogenous fixed costs associated with R&D investment. In a mixed model of vertical and horizontal product differentiation, we illustrate the theoretical lower bounds to market concentration implied by an endogenous fixed cost (EFC) model and derive the theoretical lower bound to R&D concentration from the same model. 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. For the largest-sized markets in GM corn and cotton seed, single firm concentration ratios range from approximately .35 to .44 whereas three firm concentration ratios are approximately .78 to .82. The concentration ratios for GM soybean seeds are significantly lower relative to corn and cotton, despite greater levels of product homogeneity in soybeans. Moreover, adjusting for firm consolidation via mergers and acquisitions does not significantly change the lower bound estimations for the largest-sized markets in corn or cotton for either one or three firm concentration, but does increase the predicted lower bound for GM soybean seed significantly. These results imply that concerns of concentration of intellectual property resulting from mergers and acquisitions in agricultural biotechnology are more important for some crop types relative to others.

The empirical estimations imply that the agricultural biotechnology sector is characterized by endogenous fixed costs associated with R&D investments. As firms are able to increase their market shares by increasing the quality of products offered, there are incentives for firms to increase their R&D investments prior to competing in the product market. The lower bound to concentration implies that even as the acreage of GM crops planted increases, one would not expect a corresponding increase in firm entry. However, the results from the estimations for GM soybean seeds indicate that concerns for increased concentration of intellectual property arising from firm mergers and acquisitions may be justified, even though there is little evidence to support this claim from the corn and cotton seed markets.

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 an anticompetitive 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.

References

Berry, S. and J. Waldfogel. 2003. "Product Quality and Market Size." NBER Working Paper 9675.

Chen, Q., A. Naseem, and C. Pray. 2004. "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.

Dick, A. 2007. "Market Size, Service Quality, and Competition in Banking." *Journal of Money, Credit, and Banking* 39(1): 49-81.

Dillon, M. and K. Hubbard. 2010. "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.

Ellickson, P. 2007. "Does Sutton Apply to Supermarkets?" *RAND Journal of Economics* 38(1): 43-59.

Fernandez-Cornejo, J. and M. Caswell. 2006. *First Decade of Genetically Engineered Crops in the United States*. Economic Information Bulletin 11. U.S. Department of Agriculture, Economic Research Service.

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

Fulton, M. and K. Giannakos. 2001. "Agricultural Biotechnology and Industry Structure." *AgBioForum* 4(2): 137-151.

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

Giorgetti, M. 2003. "Lower Bound Estimation – Quantile Regression and Simplex Method: An Application to Italian Manufacturing Sectors." *The Journal of Industrial Economics* 51(1): 113-120.

Goodhue, R., G. Rausser, S. Scotchmer, and L. Simon. 2002. "Biotechnology, Intellectual Property and Value Differentiation in Agriculture." CUDARE Working Paper 901R, Dept. of Agriculture and Natural Resources, UCB.

Graff, G., G. Rausser, and A. Small. 2003. "Agricultural Biotechnology's Complementary Intellectual Assets." *Review of Economics and Statistics* 85(2): 349-363.

Harhoff, D., P. Régibeau, and K. Rockett. 2001. "Some Simple Economics of GM Food." *Economic Policy* 33(16): 263-299.

Howell, V., D. Older Alguilar, J. Gleklen, and K. Letzler. 2009. "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.

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

Johnson, S.R. and T.A. Melkonyan. 2003. "Strategic Behavior and Consolidation in the Agricultural Biotechnology Industry." *American Journal of Agricultural Economics* 85(1): 216-233.

Kalaitzanodakes, N. and B. Bjornson. 1997. "Vertical and Horizontal Coordination in the Agro-biotechnology Industry: Evidence and Implications." *Journal of Agricultural and Applied Economics* 29(1): 129-139.

Lavoie, B.F. 2004. "Real Options and the Management of R&D Investment: An Analysis of Comparative Advantage, Market Structure, and Industry Dynamics in Biotechnology." Unpublished PhD Dissertation, The Ohio State University.

Marco, A. and G. Rausser. 2008. "The Role of Patent Rights in Mergers: Consolidation in Plant Biotechnology." *American Journal of Agricultural Economics* 90(1): 133-151.

Marin, P. and G. Siotis. 2007. "Innovation and Market Structure: An Empirical Evaluation of the 'Bounds Approach' in the Chemical Industry." *The Journal of Industrial Economics* 55(1): 93-111.

Moschini, G. 2010. "Competition Issues in the Seed Industry and the Role of Intellectual Property." Working Paper 10015, Iowa State University.

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

Neuman, W. 2010. "Justice Dept. Tells Farmers It Will Press Agricultural Industry on Antitrust." *The New York Times*, March 13, pp. 7.

Oehmke, J., C. Wolf, and K. Raper. 2005. "On Cyclical Industry Evolution in Agricultural Biotechnology R&D." *Journal of Agricultural & Food Industrial Organization* 3(2): 1107-1131.

Schimmelpfennig, D., C. Pray, and M. Brennan. 2004. "The Impact of Seed Industry Concentration on Innovation: A Study of U.S. Biotech Market Leaders." *Agricultural Economics* 30(2): 157-167.

Sheldon, I. 2008. "The Biotechnology Sector: "Bounds" to Market Structure". Selected Paper for presentation at the American Agricultural Economics Association Annual Meeting, Orlando, FL: July 27-29, 2008.

Smith, R. 1985. "Maximum Likelihood Estimation in a Class of Nonregular Cases." *Biometrika* 72(1): 67-90.

—. 1994. "Nonregular Regression." *Biometrika* 81(1): 173-183.

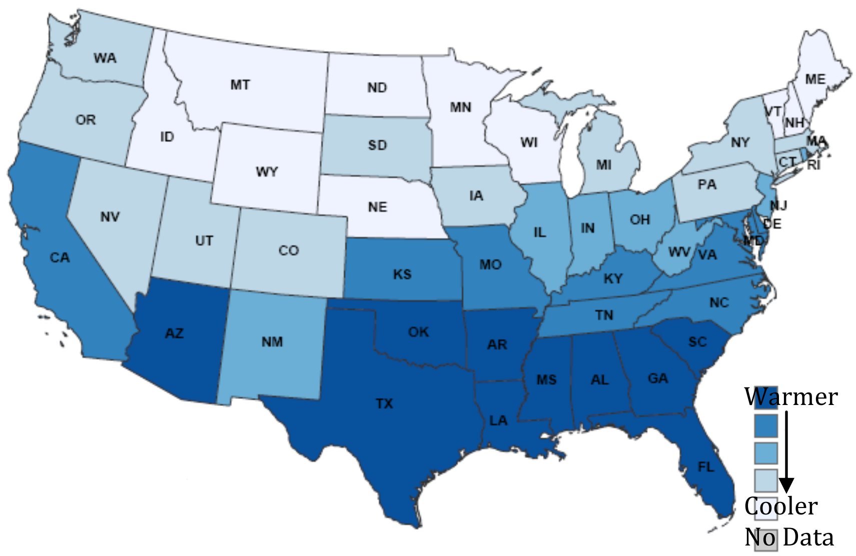
Stiegert, K., G. Shi, and J.P. Chavas. 2011. "Spatial Pricing of Genetically Modified Hybrid Corn Seeds." In C. Carter, G.C. Moschini, and I. Sheldon, eds., *Genetically Modified Food and Global Welfare: Frontiers of Economics and Globalization,* Volume 10. New Milford, CT: Emerald Books.

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

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

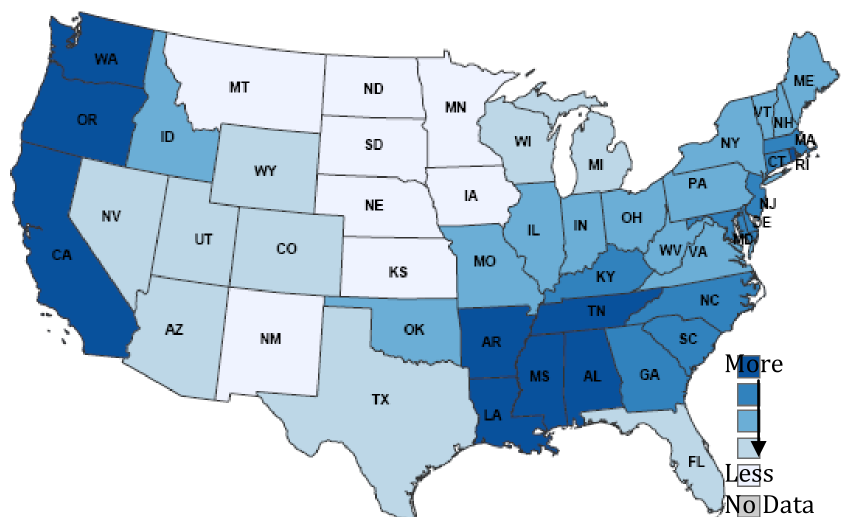
Appendix A: (Sub-)Market Analysis for GM Crops

Submarket Analysis: State-Level Climate



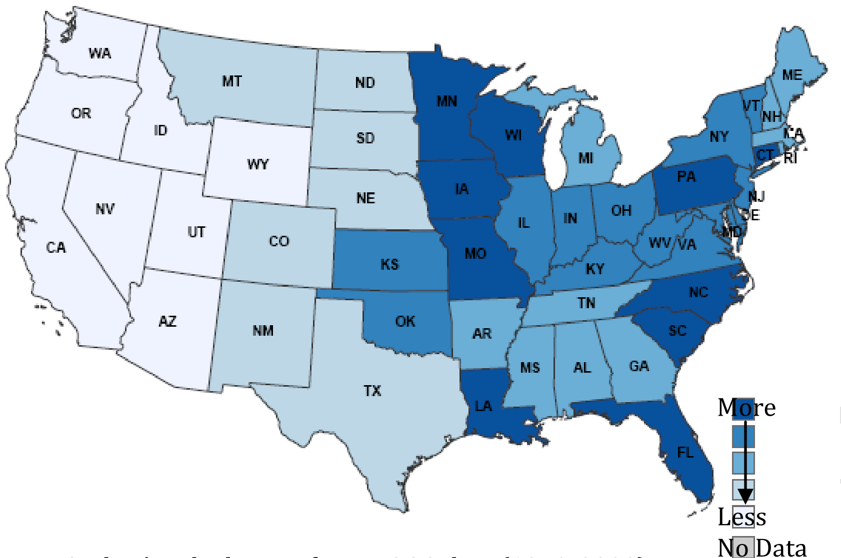
Source: Author's calculations from NOAA data (1970-2000).

Figure 16: Average Monthly Temperatures Factor Analysis



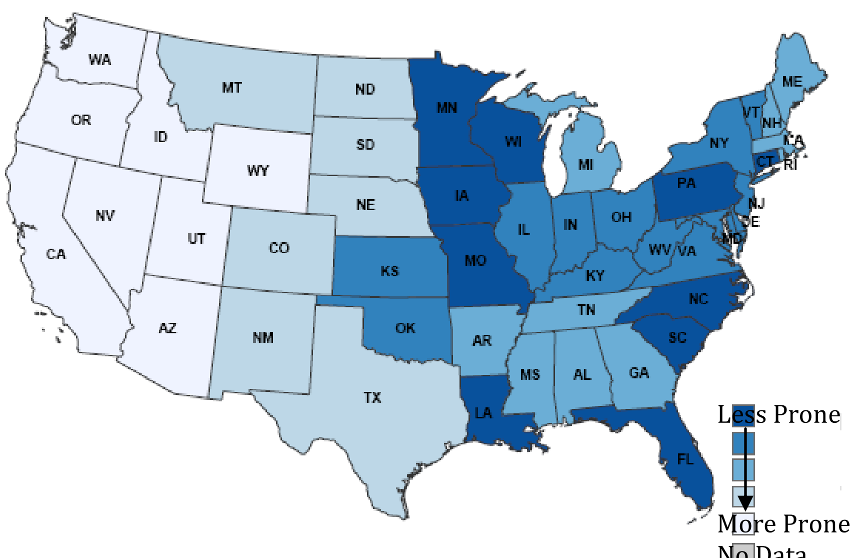
Source: Author's calculations from NOAA data (1970-2000).

Figure 17: Average Monthly Precipitation Factor Analysis (1)



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

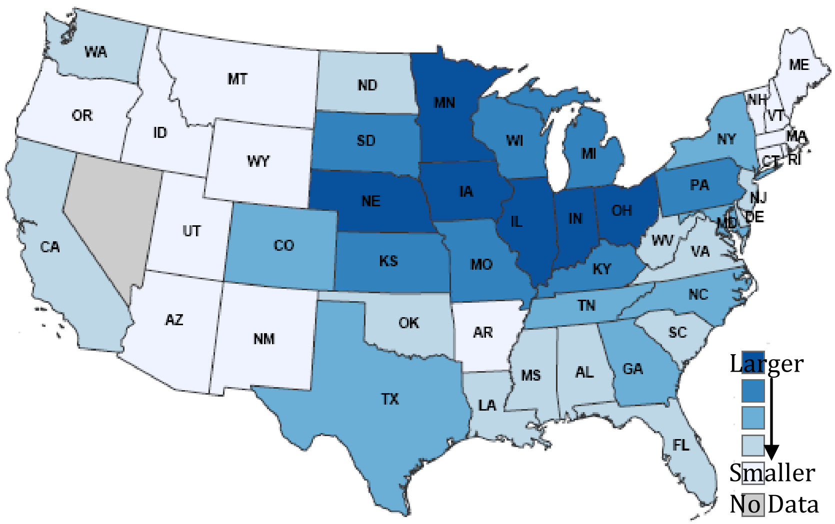
Figure 18: Average Monthly Precipitation Factor Analysis (2)



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

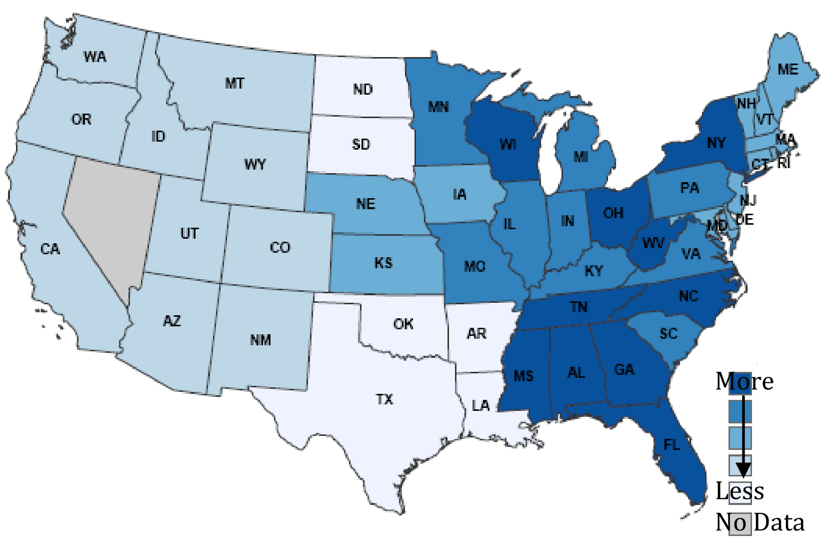
Figure 19: Average Monthly Drought Likelihood Factor Analysis

Submarket Analysis: Corn



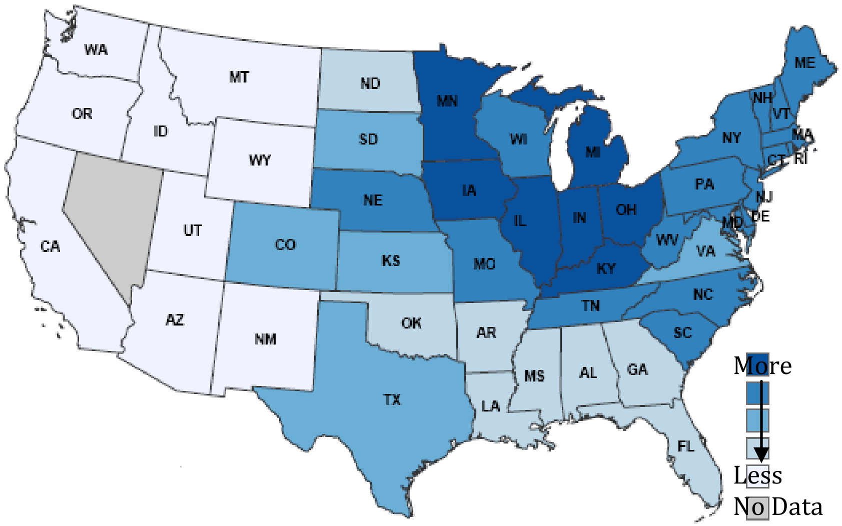
Source: Author's calculations from Census of Agriculture (1987, 1992).

Figure 20: Corn Seed Market Size Factor Analysis



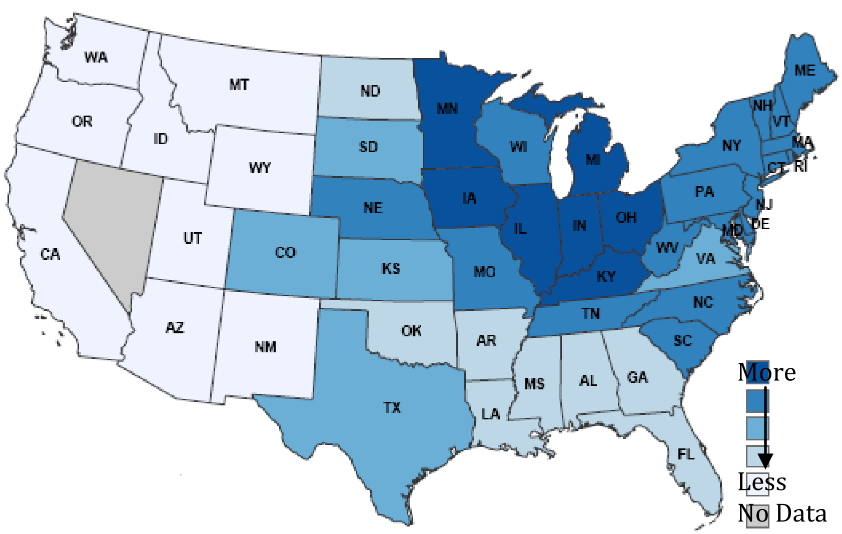
Source: Author's calculations from Agricultural Chemical Usage (1990-1995).

Figure 21: Percentage of Planted Corn Acres Treated with Fertilizer



Source: Author’s calculations from Agricultural Chemical Usage (1990-1995).

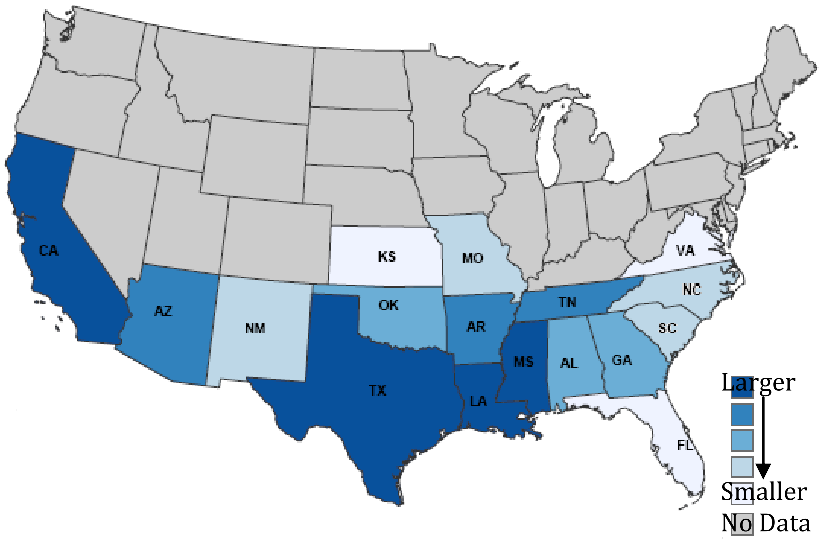
Figure 22: Percentage of Planted Corn Acres Treated with Herbicide



Source: Author’s calculations from Agricultural Chemical Usage (1990-1995).

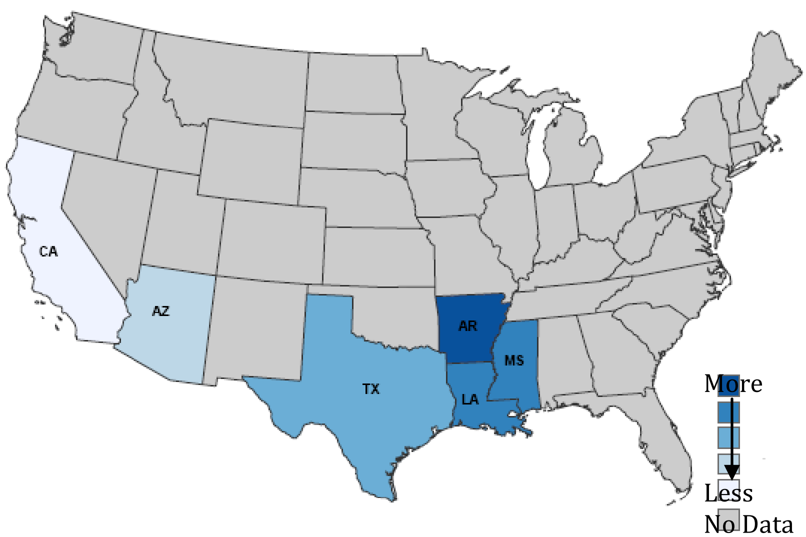
Figure 23: Percentage of Planted Corn Acres Treated with Insecticide

Submarket Analysis: Cotton



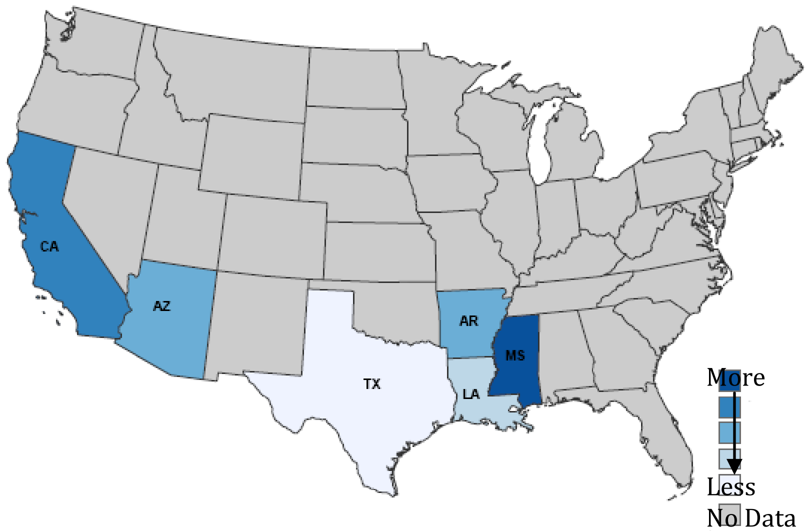
Source: Author's calculations from Census of Agriculture (1987, 1992).

Figure 24: Cotton Seed Market Size Factor Analysis



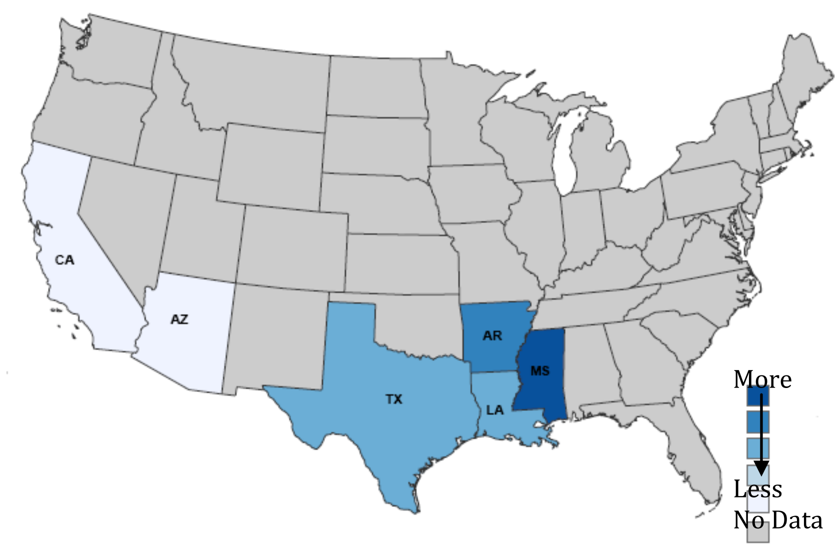
Source: Author's calculations from Agricultural Chemical Usage (1990-1995).

Figure 25: Percentage of Planted Cotton Acres Treated with Fertilizer (1)



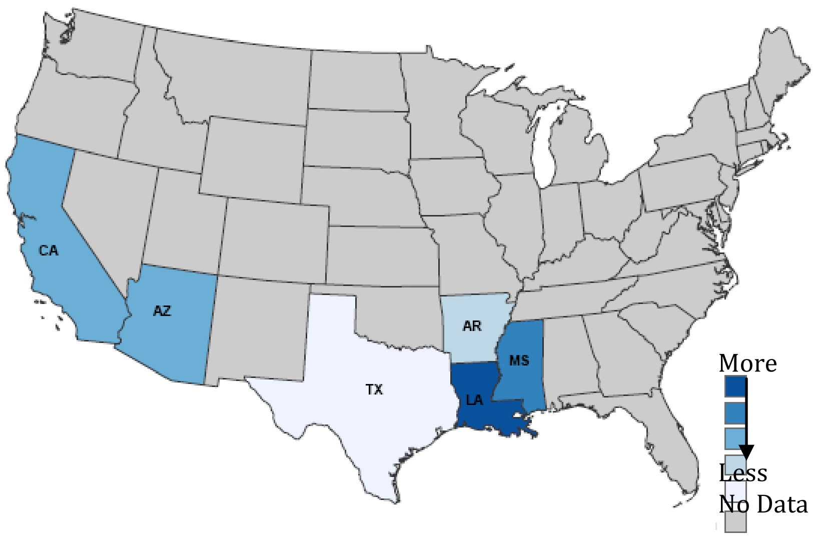
Source: Author's calculations from Agricultural Chemical Usage (1990-1995).

Figure 26: Percentage of Planted Cotton Acres Treated with Fertilizer (2)



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

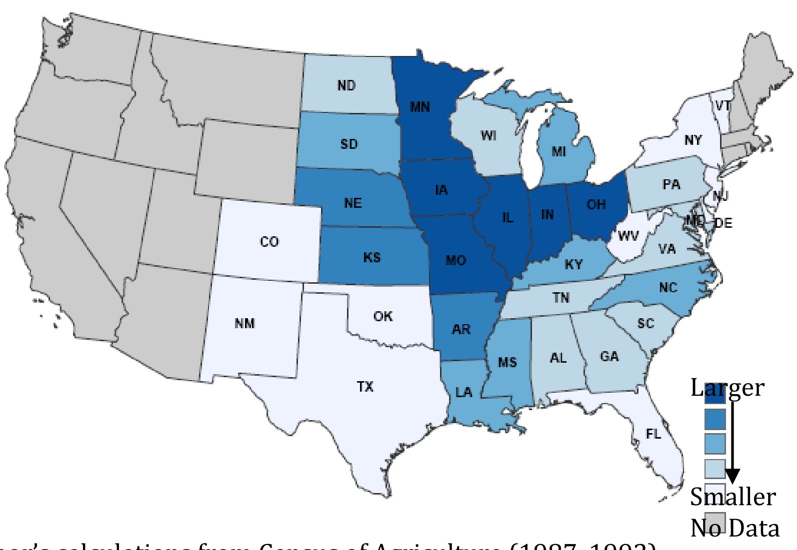
Figure 27: Percentage of Planted Cotton Acres Treated with Herbicide



Source: Author’s calculations from Agricultural Chemical Usage (1990-1995).

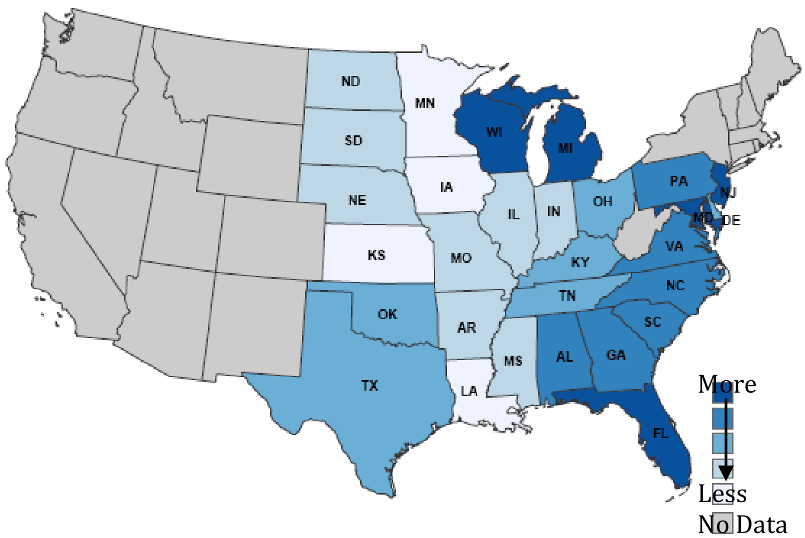
Figure 28: Percentage of Planted Cotton Acres Treated with Insecticide

Submarket Analysis: Soybean



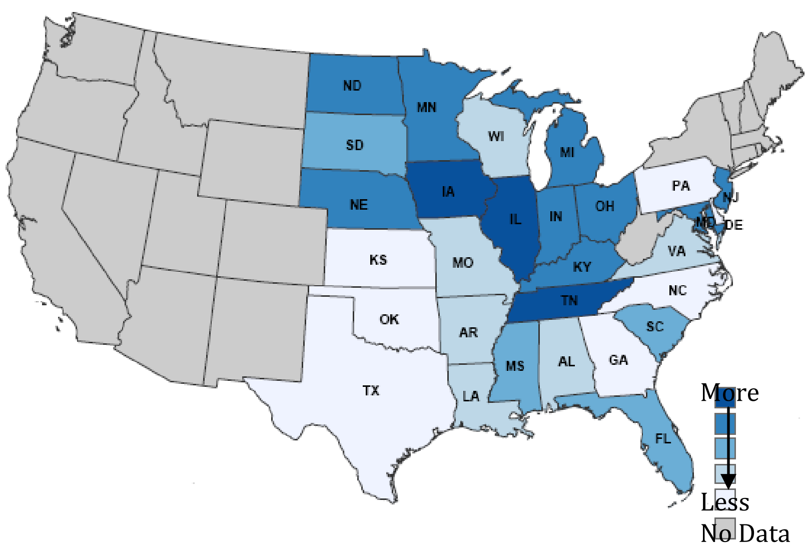
Source: Author’s calculations from Census of Agriculture (1987, 1992).

Figure 29: Soybean Seed Market Size Factor Analysis



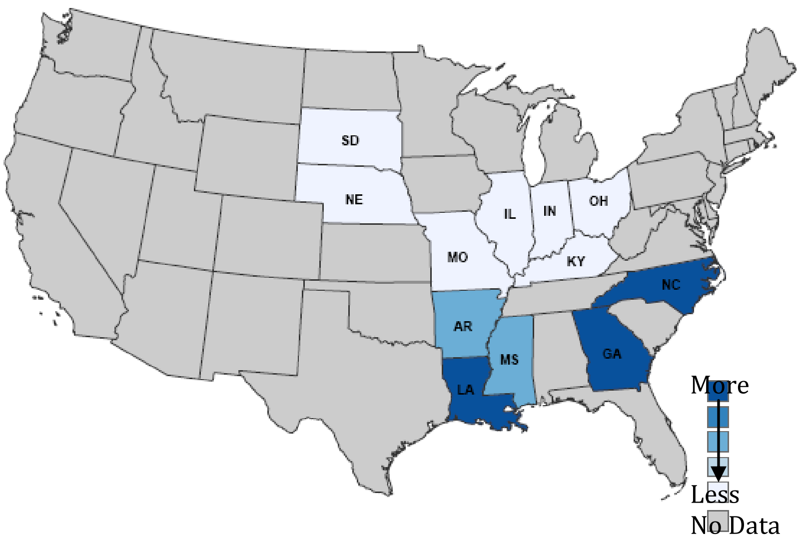
Source: Author's calculations from Agricultural Chemical Usage (1990-1995).

Figure 30: Percentage of Planted Soybean Acres Treated with Fertilizer



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

Figure 31: Percentage of Planted Soybean Acres Treated with Herbicide



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

Figure 32: Percentage of Planted Soybean Acres Treated with Insecticide