R&D Concentration in Soybean and Cotton Markets

Benjamin C. Anderson
Assistant Professor
Department of Economics
Colgate University
bcandeson@colgate.edu

Ian M. Sheldon
Professor and Andersons Chair of Agricultural Marketing, Trade, and Policy
Department of Agricultural, Environmental, and Development Economics
The Ohio State University
sheldon.1@osu.edu

Selected Paper prepared for presentation at the 2018 Agricultural & Applied Economics Association Annual Meeting, Washington, D.C., August 5-August 7

R&D Concentration in Soybean and Cotton Markets

Benjamin C. Anderson
Assistant Professor
Department of Economics
Colgate University
bcandeson@colgate.edu

Ian M. Sheldon
Professor and Andersons Chair of Agricultural Marketing, Trade, and Policy
Department of Agricultural, Environmental, and Development Economics
The Ohio State University
sheldon.1@osu.edu

PRELIMINARY AND INCOMPLETE – PLEASE DO NOT CITE WITHOUT PERMISSION

May 2018

Abstract

Using data on field trial applications, we estimate the lower bounds to concentration in research and development (R&D) activity for genetically-modified (GM) cotton and soybean seed markets. We find that both crop types exhibit endogenous costs of entry implying that firms respond to increases in market with escalations of R&D investment to improve product quality rather than permit additional firm entry. The implications of these results are that as markets for GM crop varieties become large, market concentration ratios will remain bounded away from perfectly competitive levels. In subsequent analyses, we adjust the measures of R&D concentration according to merger and acquisition (M&A) activity. We fail to find a significant impact of M&A upon increasing the concentration of intellectual property assets in GM cotton seed markets, but find that M&A significantly increased R&D concentration in smaller and medium-sized GM soybean seed markets.

Keywords: *R&D, market structure, genetically modified, cotton, soybean*

JEL Codes: L22, Q16

Acknowledgments:
This project was supported by the Agricultural Food Research Initiative of the National Institute of Food and Agriculture, USDA, Grant #2008-35400-18704. Previous versions of this article have been circulated with the title "R&D Concentration under Endogenous Fixed Costs: Evidence from the Agricultural Biotechnology Industry".

I. Introduction

The adoption of genetically-modified (GM) varieties of soybean and cotton seed has become nearly ubiquitous in the US, accounting for over 90 percent of acres planted in each crop type since their introduction in 1994 and 1995, respectively. (Wechsler, 2017) The rapid expansion of the market for GM crop varieties, coupled with frequent merger and acquisition (M&A) activity, has prompted concerns of market concentration in both the product market and in the level of innovative activity. (Moscini, 2010; Maisashvili, et al., 2016; MacDonald, 2017) Using data on field trial applications (FTA) for GM soybean and cotton varieties, we estimate the level of concentration in research and development (R&D) activity by agricultural biotechnology firms with and without accounting for M&A activity.

Figure 1 reveals that increasing adoption of GM cotton and soybean varieties since their introduction in the mid-1990s was accompanied first by an increase in the concentration of R&D activity followed by a subsequent decrease. Concentration of R&D in GM cotton seed peaked in the early 2000s whereas concentration of R&D in GM soybean seeds peaked in the mid-2000s and remain more concentrated than R&D in cotton. Anderson and Sheldon (2017) show that when R&D investments lead to increases in product quality, the concentration in R&D activity is bounded from below. These results, which Anderson and Sheldon use to estimate the level of concentration in GM corn seed markets, imply that increases in market size yield increased R&D activity by existing firms rather than permit entry by new competitors. Consistent with Lence and Hayes (2005), the overall welfare effects of this activity depend upon product quality, which can be thought of as the cost-saving to farmers of improved GM varieties, and the nature of competition in the product market.

We expand upon Anderson and Sheldon (2017) by considering two GM crop types, soybeans and cotton. Unlike the market for GM corn seed which is dominated by the US, the international market for GM varieties of soybean and cotton is substantial with global adoption rates of 78% and 64%, respectively. (ISAAA, 2016) If the financial returns to R&D activity in GM crop varieties expand beyond the domestic market alone, then we would be overestimating the level of R&D concentration along two dimensions. First, if the market share of foreign competitors abroad is substantial then the actual level of R&D concentration would be overestimated. Second, if the relevant market size for domestic firms includes international markets the fitted lower bound to R&D concentration would be "flatter" and the theoretical lower bound for large markets would be overestimated.

Prior to the release of GM varieties for R&D purposes in the US, firms, non-profits, and government organizations must file an application with the Animal and Plant Health Inspection Service (APHIS). Using FTA from the introduction of GM varieties in the late 1980s through 2010, we construct measures of R&D concentration across agro-climatic geographic regions and across time for each of the crop varieties. We use this two dimensional variation, geographic and intertemporal, in order to estimate the lower bounds to R&D concentration for GM soybean and cotton varieties with and without adjusting for M&A activity. User cluster analysis, we partition US states into non-overlapping sections in order to allow for geographic spillovers in R&D investment between similar regions. The submarkets that we define are largely consistent with the cotton production regions identified in Larson and Meyer (1996) and the soybean production regions identified in Schaub, et al. (1988).

We use a two-step procedure to estimate the lower bounds to R&D concentration. In the first step, we fit a lower bound by solving a linear programming problem using the simplex algorithm and boot-strapping the standard errors. If increases in market size lead to increased R&D activity by existing firms, the non-zero first-stage residuals should fit a two-parameter Weibull distribution. Conversely, if market size increases are accompanied by firm entry and constant (or decreasing) concentration of R&D activity, then the residuals should fit a three-parameter Weibull distribution. These distributions can be estimated via maximum likelihood and tested using a likelihood ratio test.

Our results imply that there is evidence that both GM soybean and cotton markets are characterized by increases in R&D activity by existing firms as market size increases rather than additional entry by new competitors. Moreover, although the regional submarkets for soybeans are more concentrated than the regional submarkets for cotton, there is potential for additional concentration in soybean as the fitted lower bound to concentration remains below the theoretical bound for the largest markets. Conversely, smaller-sized submarkets for cotton are already more concentrated than the theoretical minimum such that additional concentration should be viewed skeptically. Finally, accounting for previous M&A activity does not significantly change the estimates for the markets for GM cotton, but significantly increase the level of concentration in smaller and medium-sized submarkets for GM soybeans.

II. Theoretical Justification and Empirical Model of R&D Concentration

Sutton (1991, 1998, 2007) develops a model of market structure in which market entry and advertising and/or R&D investment decisions are jointly determined. When firms can vertically differentiate their products by investing in advertising or R&D, the equilibrium number of entrants, hence the concentration of firms in the market, remains bounded away from perfectly competitive levels even as the size of the market becomes large. Entry costs are considered "endogenous" in the sense that both product quality and the number of market entrants are jointly determined via investments in advertising or R&D. This contrasts with the case in which products are sufficiently homogeneous, or in which product quality is non-increasing in advertising or R&D expenditures, such that all entrant firms offer symmetric, minimum quality. In this case, firms enter the market "exogenously" such that the number of entrants (market concentration) is strictly increasing (decreasing) as market size increases.

A. A Theoretical Model of Lower Bounds to Concentration

Sutton (1998) derives empirically testable hypotheses regarding the lower bound to market concentration and the lower bound to R&D-to-sales ratio that would be observed in each case. In particular, Sutton illustrates that the market share $C_{1,m}$ of the firm offering the highest level of quality in submarket m is bounded from below under endogenous entry costs by:

(1) $$C_{1,m} \geq \alpha(\sigma, \beta) \cdot h_m$$

where $\alpha(\sigma, \beta)$ is a parameter that depends upon the degree of product substitution σ and the elasticity of R&D costs β and h_m is a measure of product homogeneity in submarket m.

Conversely, under exogenous entry costs, the market share for all entrants is symmetric and given by:

$$(2) \qquad C_{1,m} = \frac{1}{N_m}$$

where N_m is the number of entrants when all firms invest in minimum quality. When equation (1) is binding, firms respond to an increase in the size of the submarket by escalating product quality rather than permitting entry by additional firms. Equation (2) implies that an increase in submarket size will result in entry by additional firms such that the concentration ratio is strictly decreasing. Although these two conditions are stated separately, it is important to note that a single industry may be characterized by either endogenous or exogenous entry costs depending upon the underlying parameters. Equation (1) is more likely to arise when products are more homogeneous or closer substitutes, when R&D costs are lower, and when submarket sizes are larger. The extent that R&D and/or advertising investments jointly determine product quality and the number of entrant firms can be tested empirically via cross-industry analysis (Robinson and Chiang, 1996; Sutton, 1998) or across submarkets m (Sutton, 1991; Latcovich and Smith, 2001; Dick, 2007; Ellickson, 2007; Marin and Siotis, 2007, Berry and Waldfogel, 2010; Anderson and Sheldon, 2017).

Anderson and Sheldon (2017) show that under the same conditions identified in Sutton (1998), the lower bound to concentration in R&D expenditures can also be derived and empirically tested. When R&D investments and market entry decisions are endogenous, then the firm investing in the market-leading level of quality in submarket m will have a share of R&D expenditures $R_{1,m}$ bounded from below by:

(3) $$R_{1,m} \geq \left[\alpha^2(\sigma,\beta)\cdot h_m^2 - \alpha(\sigma,\beta)\cdot h_m\left(\frac{F_0}{S_m y_m}\right)\right]$$

where F_0 is the fixed setup cost associated with entry, S_m is the number of consumers in submarket m, and y_m is the industry sales revenue in submarket m. Conversely, the upper bound on R&D concentration is bounded from above by:

(4) $$R_{1,m} \leq \frac{1}{N_m}$$

Since $\alpha(\cdot) \in [0,1]$ and $h_m \in [0,1]$, the lower bound to concentration in R&D expenditure when firms make quality-enhancing R&D investments with entry is less than the lower bound to market concentration. Moreover, as the size of the submarket, in terms of the number of consumers or the total industry revenue, increases, the lower bound to R&D concentration increases. This implies that larger submarkets are more likely to be concentrated compared with relatively smaller submarkets. For markets in which entry costs are exogenous, the level of market concentration forms an upper bound on the level of R&D concentration.

B. Empirical Specification

The lower bounds to R&D concentration, reflected in equations (3) and (4), can be estimated using maximum likelihood when the concentration ratio is characterized by a Weibull distribution (Anderson and Sheldon, 2017). We follow the empirical estimation strategy developed in Sutton (1991) based upon Smith (1985; 1994) and the simplex methodology of Giorgetti (2003). In order to derive the empirically testable equations, the R&D concentration ratio must be transformed such that the predicted concentration measures lie between 0 and 1. We first monotonically shift $R_{1,m}$ by -0.0001 to address submarkets with only a single entrant and then transform the concentration measure according to:

(5)
$$\tilde{R}_{1,m} = \ln\left(\frac{R_{1,m}}{1 - R_{1,m}}\right)$$

In order to estimate the lower bounds to R&D concentration, the transformed concentration measure for each submarket m is normalized by the degree of product homogeneity in the submarket such that the functional form for the estimation is:

(6)
$$\frac{\tilde{R}_{1,m}}{h_m^2} = \theta_0 - \theta_1\left(\frac{1}{h_m \ln(S_m y_m / F_0)}\right) + \varepsilon_m$$

Estimates of θ_0, the theoretical lower bound to market concentration for large markets, and θ_1, the slope parameter for changes in the lower bound as market size changes, can be obtained via a linear programming problem such that the residuals ε_m are non-negative. The fitted residuals follow a Weibull distribution such that:

(7)
$$F(\varepsilon) = 1 - \exp\left[-\left(\frac{\varepsilon - \mu}{\delta}\right)^{\gamma}\right], \gamma > 0, \delta > 0$$

where $\varepsilon \geq \mu$. The Weibull distribution is characterized by three parameters (μ, δ, γ) which reflect the "shift", "scale", and "shape" of the distribution. The shift parameter μ represents the degree of horizontal shift of the distribution such that when $\mu = 0$ corresponds to a two-parameter Weibull distribution. The scale parameter δ reflects the dispersion of the Weibull distribution and the shape parameter γ captures the degree of clustering around the lower bound.

The estimation of the lower bound to R&D concentration involves a two-step procedure. First, we obtain consistent estimates of the lower bound parameters $\hat{\theta}_0$ and $\hat{\theta}_1$

by solving a linear programming problem using the simplex algorithm under the constraint that the model residuals are non-negative such that:

$$\text{(8)} \qquad \begin{aligned} &\min_{\{\theta_0,\theta_1\}} \sum_{m=1}^{M} \left[\frac{\tilde{R}_{1,m}}{h_m^2} - \left(\theta_0 - \theta_1 \left(\frac{1}{h_m \ln(S_m y_m / F_0)} \right) \right) \right] \\ &\textit{subject to } \frac{\tilde{R}_{1,m}}{h_m^2} \geq \theta_0 - \theta_1 \left(\frac{1}{h_m \ln(S_m y_m / F_0)} \right), \forall m \end{aligned}$$

with standard errors that can be calculated via bootstrapping. Since there are two first-stage parameters, there will be $M-2$ positive fitted residuals $\hat{\varepsilon}_m$ from the first stage. These residuals can be used to estimate the Weibull distribution parameters (μ, δ, γ) via the maximization of the log pseudo-likelihood function:

$$\text{(9)} \qquad \max_{\{\mu,\delta,\gamma\}} \sum_{m=1}^{M-2} \ln \left[\left(\frac{\gamma}{\delta} \right) \left(\frac{\hat{\varepsilon}_m - \mu}{\delta} \right)^{\gamma - 1} \exp \left[- \left(\frac{\hat{\varepsilon}_m - \mu}{\delta} \right)^{\gamma} \right] \right]$$

with standard errors that can be estimated according to the asymptotic distributions defined by Smith (1994). The shift parameter estimate $\hat{\mu}$ can be used to test for the validity of the three-parameter Weibull distribution ($\hat{\mu} > 0$) against the restricted, two-parameter Weibull distribution ($\hat{\mu} = 0$). Failure to reject the three-parameter Weibull distribution implies that we are unable to reject that R&D expenditures are exogenous since the transformed measures of R&D concentration are shifted away from the lower bound.

III. Data and Descriptive Statistics

The lower bound to R&D concentration in soybean and cotton seed markets can be estimated according to equation (8) using data for each crop type at the submarket level. We exploit variation in the adoption and prevalence of GM crop varieties across two dimensions: (i) suitability of GM traits to agro-climatic conditions that vary geographically, and (ii) intertemporal variation in the adoption and expansion of GM crop varieties across geographic submarkets. An estimation of the lower bound to R&D concentration in these seed markets requires data at the firm level on R&D investment $R_{1,m}$ for each crop type and every submarket. We aggregate state-level data on field trial applications (FTA) of GM crops into geographically distinct submarkets as a proxy for firm-level R&D investment. In addition to the firm-level data used to calculate the degree of R&D concentration, the lower bounds to R&D concentration also depend upon industry-level data on submarket size $S_m y_m$, the degree of product homogeneity in the submarket h_m, and the minimum setup costs F_0 that a firm must incur in order to enter a submarket.

A. Geographic Submarket Cluster Analysis

In order to estimate the lower bounds to R&D concentration in a single industry, we first must identify distinct submarkets. Previous industry-level analyses of lower bounds to concentration have largely focused on retail industries which can be separated spatially. These include examinations of retail banking (Dick, 2007), supermarkets and barbers/beauty salons (Ellickson, 2007), and newspapers and restaurants (Berry and Waldfogel, 2010). Unlike a retail environment in which firms incur advertising expenditures in each submarket, R&D investment in GM traits face a greater potential for spillovers across submarkets. The potential for spillovers across submarkets rules out the possibility of using patent applications as a proxy for R&D activity since these occur at the national level and are equally applicable to all submarkets.

In order to identify the relevant subnational geographic markets for seed varieties, we make a critical identifying assumption that R&D investments in GM seed varieties are recouped within a particular geographic submarket only. Specifically, we assume that if a firm wishes to market its existing GM seed in a different geographic submarket, then it first must test those varieties in the submarket that it wishes to enter. This assumption motivates us to characterize geographic submarkets for soybean and cotton seed according to observable agricultural and climatic differences.

Cluster analysis permits us to partition states into regional clusters following a "natural structure" of observable agricultural and climate characteristics. We assume a "prototype-based" framework such that every state in a cluster is more similar to a prototype state for that submarket than it is to every other submarket's prototype state. We utilize a K-means approach by defining the number of K submarket clusters for each crop type and minimizing the Euclidean distance between each state and the centroid of the cluster. For robustness, we consider alternate K clusters for each crop type as well as minimizing the absolute distance function.[1] The results of the cluster analysis are reported in Table 1 along with the corresponding market shares of US production and number of field trial applications for each submarket.

B. Measuring R&D Concentration

The ideal data for measuring R&D concentration in GM seed markets would be R&D expenditures for each product line in each submarket for every active firm. Although this level of detail on R&D expenditures is unavailable for GM crop varieties, there is publicly available data on field trial applications (FTA) that capture an intermediate stage of the R&D process. The Biotechnology Regulatory Services (BRS), a division of the Animal and

[1] For additional information regarding the cluster analysis, please refer to the appendix.

Plant Health Inspection Services (APHIS), mandates that all importation, interstate movement, and release of GM organisms are reported by firms and organizations. BRS publishes this database of permits, notifications, and petition applications which includes information on the applicant institution, the status of the application, the plant (or "article") type, the dates 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. Our data covers 1985 through 2010 consisting of 33,440 permits or notifications of release across all crop types. We restrict the sample to include only for-profit firms and to limit ourselves to applications pertaining to the release of GM soybean and cotton varieties.

Figure 2 plots the annual number of field trial applications, the number of firms that file an application for a field trial each year, and the average number of applications per firm by year for both GM cotton and soybean varieties. Although the number of individual firms peaked in the early 1990s for both cotton and soybeans, the total number of applications did not peak until the early 2000s for cotton and the late 2000s for soybean even though the number of firms had fallen from previous highs. The number of applications per firm reflect this increased intensity of R&D with the intensity of cotton research peaking in the late 1990s and early 2000s and the intensity of soybean applications peaking in the mid to late 2000s.

We aggregate the number of FTA for each firm at the geographic submarket level across five-year intervals in order to derive a measure of R&D concentration that varies across submarkets and across time. We account for potential geographic spillovers by aggregating applications up to the submarket level consisting of states with similar agricultural and climatic characteristics. We aggregate across multiple years to account for the longer-term nature of the research and development process in which year-to-year fluctuations are secondary to longer-run trends.

C. Industry-level Data on Market Size, Product Homogeneity, and Setup Costs

In order to estimate the lower bounds to R&D concentration, we still require submarket size, a measure of product homogeneity at the submarket level, and minimum R&D setup costs. Our primary measure of submarket size is a proxy for total industry sales that we construct using annual data from the June Agriculture Survey and the Agricultural Resource Management Surveys (ARMS). We obtain acreage reports on the total acres planted and harvested at the crop level from acreage reports from June Agriculture Surveys for each state and aggregate within submarkets. The Economic Research Service (ERS) computes yearly seed costs for each crop type based upon ARMS data. After adjusting for inflation, we multiply the annual seed costs by the total acres planted to obtain our proxy for industry sales at the submarket level.

As a robustness check, we consider a definition of industry sales at the submarket level for GM seed varieties only. We combine estimates on adoption of GM seed varieties by Fernandez-Cornejo and McBride (2002) for 1996-1999 with estimates provided by the June Agriculture Surveys for 2000-2010 in order to obtain a proxy for submarket-level industry sales of GM crop varieties.[2] These rates of adoption are also used to construct the degree of product homogeneity for each crop type at the submarket level. By definition, the

[2] Although estimates of the adoption rates are available for a subsample of states only, the National Agricultural Statistics Service (NASS) reports that the sample states cover 87-90% of all soybean acres planted and 81-93% of all upland cotton acres planted. For states without an estimate for adoption rates, overall US adoption estimates are used to compute the size of the GM market.

product homogeneity index is meant to capture the percentage of industry sales of the largest product group. We consider product groups as broadly defined: conventionally-bred varieties, insect resistant (IR) varieties, herbicide tolerant (HT) varieties, and "stacked" varieties consisting of both IR and HT traits. The product homogeneity index is calculated as the percentage of acres planted with the largest product group.[3]

The final data required to estimate the lower bounds to R&D concentration is the minimum setup cost associated with entry into the product market. For each crop type, we obtain an estimate of minimum setup cost by summing the total number of public "scientist years" (SY), as reported by the State Agricultural Experiment Stations (SAES) and the Agriculture Research Service (ARS), and dividing this sum by the total number of reported projects to obtain an average SY per crop. Using data from Frey (1996) and Traxler, et al. (2006), we multiple the average SY by the private industry cost per SY ($148,000) and adjust for inflation.[4]

IV. Empirical Results and Discussion

The unadjusted measures of R&D concentration, both for the market-leading firm in each submarket and for the largest four firms in each submarket, are plotted against the size of each submarket in Figure 3 for cotton seeds and Figure 4 for soybean seeds. The raw data reveals a considerable amount of concentration across submarkets and across time with the four-firm concentration ratios in cotton seed exceeding 0.75 in every submarket and

[3] In a robustness check, we also consider the possibility that measurement error biases the results towards finding endogenous R&D investments by assuming perfectly homogeneous products.

[4] 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)." We also consider a robustness check using the public sector cost per SY ($296,750) such that minimum entry costs are higher.

the four-firm concentration ratios in soybean seed exceeding 0.60 in every submarket. The single-firm R&D concentration ratios for both GM cotton and soybean seeds also appear to be non-decreasing in market size.

In order to estimate the lower bounds to R&D concentration, the raw data presented in Figures 3 and 4 is transformed according to equation (5) and the lower bounds are estimated controlling for the degree of product homogeneity in each submarket. The baseline, two-stage estimation results are reported in Table 2 and the estimated lower bounds are illustrated in Figures 5 and 6 for cotton and soybean seeds, respectively. Direct interpretation of the coefficients on the lower bound estimations can be difficult due to the logit transformation of the measure of R&D concentration. However, the first-stage intercept estimate $\hat{\theta}_0$, when adjusted for product homogeneity and transformed by the inverse logit function, is equivalent to the theoretical lower bound to R&D concentration as market size becomes large. The coefficient on adjusted market size $\hat{\theta}_1$, after adjusted for product homogeneity and fixed setup costs, informs us as to whether R&D concentration is increasing (a negative parameter), decreasing (a positive parameter), or independent (insignificant parameter) in the submarket size.

The first-stage estimates reveal that there exists a lower bound to R&D concentration that does not converge to zero as the market size becomes large. Moreover, the increasing lower bound is consistent with an industry in which increasing in market size are accompanied by escalations in R&D, in order to improve product quality, rather than permit entry by additional firms. The results from Table 2 imply that the largest firm, in an infinitely-sized submarket, would account for 47.3% of the R&D in cotton seed and 78.6% of the R&D in soybean seeds. Although R&D in the largest submarket in the latest time period (2006-2010) is already substantially concentrated in cotton, with a fitted R&D share for the largest firm of 35.9%, it is much less concentrated in soybeans at 20.8% especially compared with the theoretical predictions. This reveals that substantial consolidation of R&D activity in soybean seeds following the end of the sample in 2010 could still be expected.

From the likelihood ratio tests of the second-stage results, we fail to reject the hypothesis that the first-stage residuals fit a two-parameter Weibull distribution for both cotton and soybeans seeds. That is, we fail to reject the hypothesis that $\mu = 0$ such that there is no horizontal shift of the distribution that would be consistent with a poor fit of residuals that would arise under exogenous R&D costs. The estimated shape parameter $\hat{\gamma}$ is remarkably similar for cotton and soybean seeds and indicates a fair amount of clustering of observations around the lower bound to R&D concentration. Parameter estimates that are less than two confirm the appropriateness of Smith's (1985, 1994) two-step procedure. The estimated scale parameter $\hat{\delta}$ is also consistent with a relatively narrow dispersion of first-stage residuals.

A. Robustness Checks

In order to test the validity of our estimations of the lower bounds to R&D concentration in cotton and soybean seed markets, we consider four robustness checks. First, we consider an alternate definition of submarket size based solely upon an estimate of the number of acres planted with GM varieties only. We subsequently explore the role of the product homogeneity index upon our results by assuming homogenous products which would bias our results towards finding R&D costs to be exogenous. The third robustness check considers an alternate definition of minimum setup costs based upon public sector SY which exceed private sector SY. The greater minimum setup costs should again bias our estimations towards failing to find exogenous R&D costs. Finally, we consider an alternate functional form for market size as proposed by Dick (2007) which allows for the lower bound to R&D concentration to change non-monotonically in market size.

The robustness checks, reported in Table 3, generally support our findings of endogenous R&D investments in GM cotton and soybean seed markets. The results reported in columns labeled "GM Market Size" differ from our baseline estimations in two dimensions. We limit the submarket size to only those acres planted with GM varieties and also restrict our sample to observations between 1996 and 2010 since commercially available GM varieties were not available between 1991 and 1995. The results from these estimations confirm our baseline estimations of endogenous R&D investments in both sign and magnitude. The theoretical lower bounds to concentration implied by the first-stage estimates increase for GM cotton seeds from 0.473 to 0.600 as well as for GM soybean seeds from 0.786 to 0.953 with both fitted lower bounds increasing more rapidly under the alternate submarket definition.

In the second robustness check, we explore the measurement of the product homogeneity index upon our estimations of endogenous R&D investments. These results, reported in the "Homogeneity" columns confirm the importance of firms being able to differentiate their products in order to capitalize upon investments in quality. When the product differentiation channel is "turned off" the estimation results are consistent with exogenous R&D investments as expected. The "Setup Costs" columns report results in which the minimum setup costs are assumed to be consistent with the public sector cost of

R&D which exceed private sector costs. Neither the theoretical nor the fitted lower bounds to R&D concentration are substantially changed for either cotton or soybean seeds when the setup costs are assumed to be higher suggesting that the measurement of minimum setup costs is not driving the endogenous R&D investment results.

Finally, we consider an alternate specification to permit a nonlinear relationship between the lower bound to R&D concentration and submarket size. These results, presented in the "Quadratic Market Size" columns in Table 3, confirm the estimates of GM soybean seeds being characterized by endogenous R&D costs. However, the parameter estimates for GM cotton seed imply that the fitted lower bound to R&D concentration is decreasing in submarket size which is consistent with exogenous R&D costs. An examination into the underlying data reveals that a single outlier, Western states in 2006-2010, is driving these estimation results. In unreported estimations eliminating this single outlier, which consists of a small-sized submarket and low levels of concentration, the market of GM cotton seed is again characterized by endogenous R&D costs.

B. Impacts of Mergers and Acquisitions

In order to examine the impact of mergers and acquisitions (M&A) upon R&D investment and concentration, we adjust the field trial application data to account for changes in ownership of intellectual property. If M&A are resulting in intellectual property assets becoming more concentrated in a small number of firms, then it is possible that the lower bound to R&D concentration increases not due to the presence of endogenous R&D costs, but rather due to this consolidation activity. We utilize company histories and Lexis-Nexis new releases in order to identify M&A activity and the effective merger date in order to construct a measure of R&D concentration accounting for ownership changes. Although completed independently, our list of changes in corporate ownership corresponds to the activity reported in Fuglie, et al. (2011).

The estimations of the lower bounds to R&D concentration adjusting for M&A activity are reported for cotton and soybean seeds in Table 4. We continue to reject the null hypothesis of exogenous R&D costs for GM cotton seed with a theoretical lower bound that is significantly different from zero and a fitted lower bound that is non-decreasing in market size. Figure 7 illustrates the theoretical and fitted lower bounds with and without the adjustment for M&A activity for GM cotton seed. Equivalence tests between the parameter estimates with and without adjusting for changes in intellectual property ownership are reported in Table 5. We fail to find evidence supporting the hypothesis that M&A activity significantly increased the lower bounds to R&D concentration in GM cotton.

Conversely, the lower bound estimation results for GM soybean seeds change significantly and we fail to reject the hypothesis of exogenous fixed costs (Table 4). These results contrast with the expected results in which M&A activity would increase R&D concentration in these markets. Table 5 reports tests for significant differences in the estimated parameters which confirm these results as the theoretical lower bound accounting for M&A activity is significantly lower and the fitted lower bound is now decreasing, rather than increasing, in submarket size. Figure 8 illustrates the impact upon the estimated lower bounds to R&D concentration in GM soybean seeds when accounting for M&A activity. Inspection of Figures 6 and 8 reveal that the significant shifts in the lower bounds to R&D concentration occur due to increased concentration in medium-sized markets with no change in the largest market sizes.

V. Conclusion

Using geographic variation in R&D activity, we analyze markets for genetically-modified cotton and soybean varieties in order to determine whether increases in market size leads to additional firm entry or if existing firms can preclude additional entry by escalating R&D in order to improve product quality. Using data on field trial applications, an intermediate stage in the R&D process for GM crops, we estimate the lower bounds to R&D concentration for cotton and soybean seeds. Our results imply that both GM cotton and soybean markets are characterized by endogenous R&D investments and are robust to alternate definitions of market size and setup costs. Merger and acquisition activity, which is argued to increase the concentration of intellectual property assets, fails to significantly change the lower bounds to R&D concentration for GM cotton but do increase concentration in medium-sized markets for GM soybeans.

Our empirical results are of interest in the face of continuing concerns regarding concentration in agricultural inputs especially in light of the recent announced mergers between Bayer and Monsanto and between Dow Chemical and DuPont. Although these mergers are not reflected in our empirical results adjusting R&D concentration for changes in intellectual property ownership, Bayer, Dow, and Monsanto were three of the five firms (adjusted for previous M&A activity) that had conducted field trials in cotton between 2006 and 2010. Moreover, Bayer and Monsanto were the two most prolific firms in terms of field trials for GM cotton such that R&D concentration in each cotton submarket would increase from between 43-53% to between 77-87%.

The situation is less severe for R&D related to GM soybean seeds. Adjusting for previous M&A activity, Monsanto was the market-leading firm in the number of field trials in every soybean submarket. However, the merger between Monsanto and Bayer would increase the unadjusted market share in each submarket by less than 2%. The largest impact in soybean markets may be felt from the merger between Dow and DuPont which previously had been the second and third largest firms, in terms of field trials, between 2006 and 2010. The market share of the second-largest firm in each submarket would subsequently increase from between 12-18% to 21-32%.

An additional merger, between Syngenta and ChemChina, reflects one of the main limitations of our current analysis. Although Syngenta has tested GM crops in the US since it was formed in 2000, and previously had tested in the US as Novartis and Ciba-Geigy, ChemChina did not test GM crops in the US during our sample period. Any implications of this merger upon GM crops and the agricultural biotechnology industry would likely be reflected in international markets including China, India, Pakistan, and Brazil for cotton and Brazil, Argentina, and China for soybean production. A key identifying assumption of our empirical model is that the R&D investment associated with field trial applications that are conducted within the US can be recouped solely from the US market. If the R&D investments from the largest submarkets are recouped across a larger international market, the results identifying R&D investments as being endogenous would be unaffected. The major caveat would be if R&D investments that are made in smaller, less concentrated markets are recouped internationally such that the predicted lower bound to R&D concentration "switches" from increasing to decreasing such that we would be unable to reject R&D costs as exogenous. Given the regulatory hurdles that exist to gain approval across different countries, such that the minimum setup costs are also likely to be higher than measured, this scenario is unlikely.

References

Anderson, B.C. and I.M. Sheldon. "R&D Concentration Under Endogenous Fixed Costs: Evidence from Genetically Modified Corn Seed." *American Journal of Agricultural Economics* 99(5), 2017: 1265-1286.

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

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

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

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

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

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

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

ISAAA. *Global Status of Commercialized Biotech/GM Crops: 2016*. ISAAA Brief No. 52, ISAAA: Cornell, NY, 2016.

Larson, J.A. and L.A. Meyer. "Chapter 1: Supply, Demand, and Prices." In E.H. Glade, Jr., L.A. Meyer, and H. Stults (eds.) *The Cotton Industry in the United States*. Agricultural Economic Report AER739, Commercial Agriculture Division, Economic Research Service, U.S. Department of Agriculture, 1996.

Latcovich, S. and H. Smith. "Pricing, Sunk Costs, and Market Structure Online: Evidence from Book Retailing." *Oxford Review of Economics* 17(2), 2001: 217-234.

Lence, S.H. and D.J. Hayes. "Genetically Modified Crops: Their Market and Welfare Impacts." *American Journal of Agricultural Economics* 87(4), 2005: 931-950.

MacDonald, J.M. "Mergers and Competition in Seed and Agricultural Chemical Markets." *Amber Waves*, Economic Research Service, US Department of Agriculture, April 3, 2017. Retrieved 2018-01-12, https://www.ers.usda.gov/amber-waves/2017/april/mergers-and-competition-in-seed-and-agricultural-chemical-markets/.

Maisashvili, A., H. Bryant, J.M. Raulston, G. Knapek, J. Outlaw, and J. Richard. "Seed Prices, Proposed Mergers and Acquisitions Among Biotech Firms" *Choices* 31(4), 2016.

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

Moscini, G. "Competition Issues in the Seed Industry and the Role of Intellectual Property." *Choice* 25(2), 2010.

Robinson, W.T. and J. Chiang. "Are Sutton's Predictions Robust?: Empirical Insights into Advertising, R&D, and Concentration." *The Journal of Industrial Economics* 44(4), 1996: 389-408.

Schaub, J., W.C. McArthur, D. Hacklander, J. Glauber, M. Leath, and H. Doty. *The U.S. Soybean Industry*. Agricultural Economic Report AER588, Economic Research Service, U.S. Department of Agriculture, 1988.

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

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

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

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

—. "Market Structure: Theory and Evidence." In M. Armstrong and R.H. Porter (eds) *Handbook of Industrial Organization* Vol. 3, 2007: 2301-2368.

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

Wechsler, S.J. "Recent Trends in GE Adoption." Economic Research Service, U.S. Department of Agriculture, 2017. Retrieved 2018-01-12, https://www.ers.usda.gov/data-products/adoption-of-genetically-engineered-crops-in-the-us/recent-trends-in-ge-adoption.aspx

Figure 1: R&D Concentration Ratios and Adoption of GM Cotton and Soybean

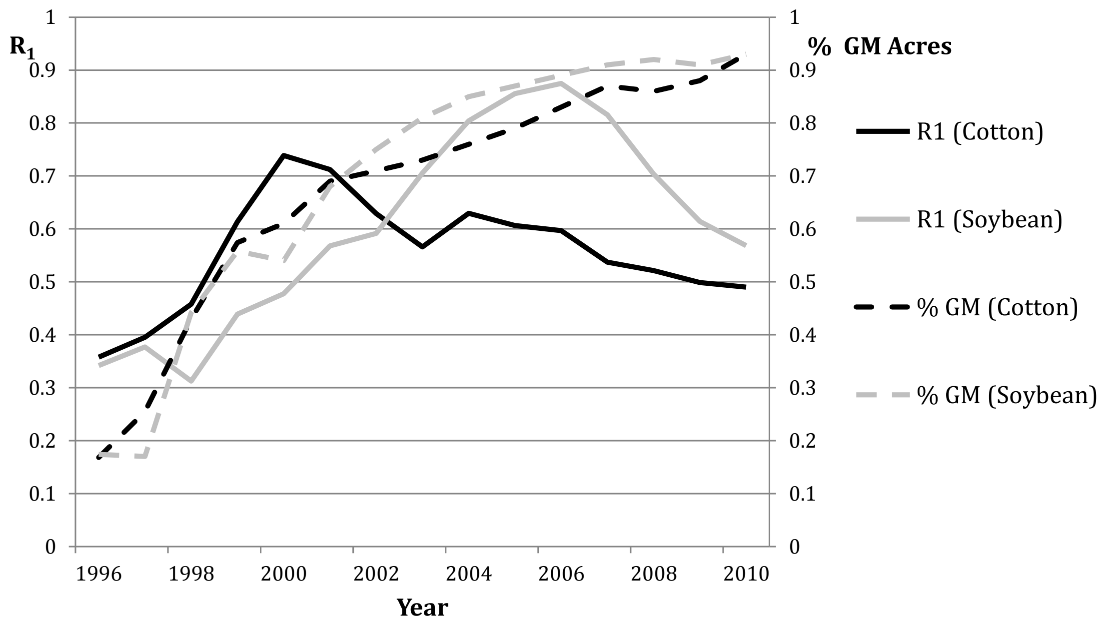


Figure 2: Field Trial Applications and Firms in GM Cotton and Soybean Seed Markets

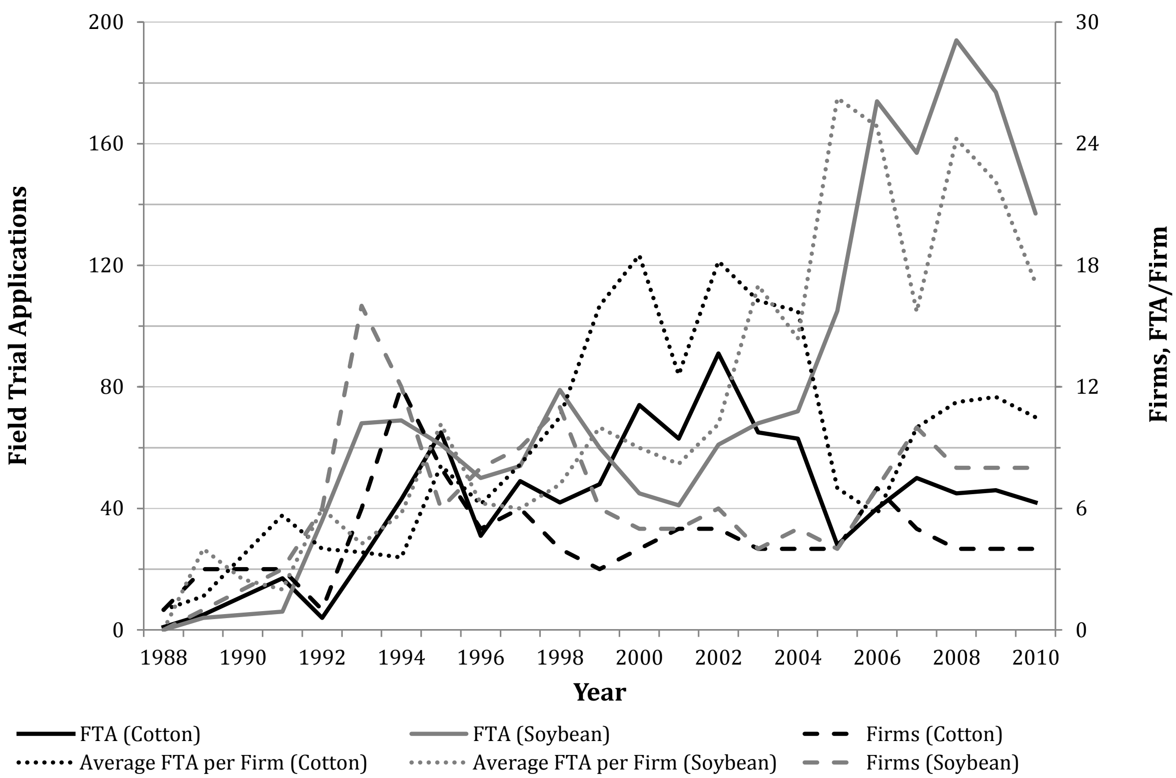


Figure 3: R&D Concentration and Market Size in GM Cotton Seed

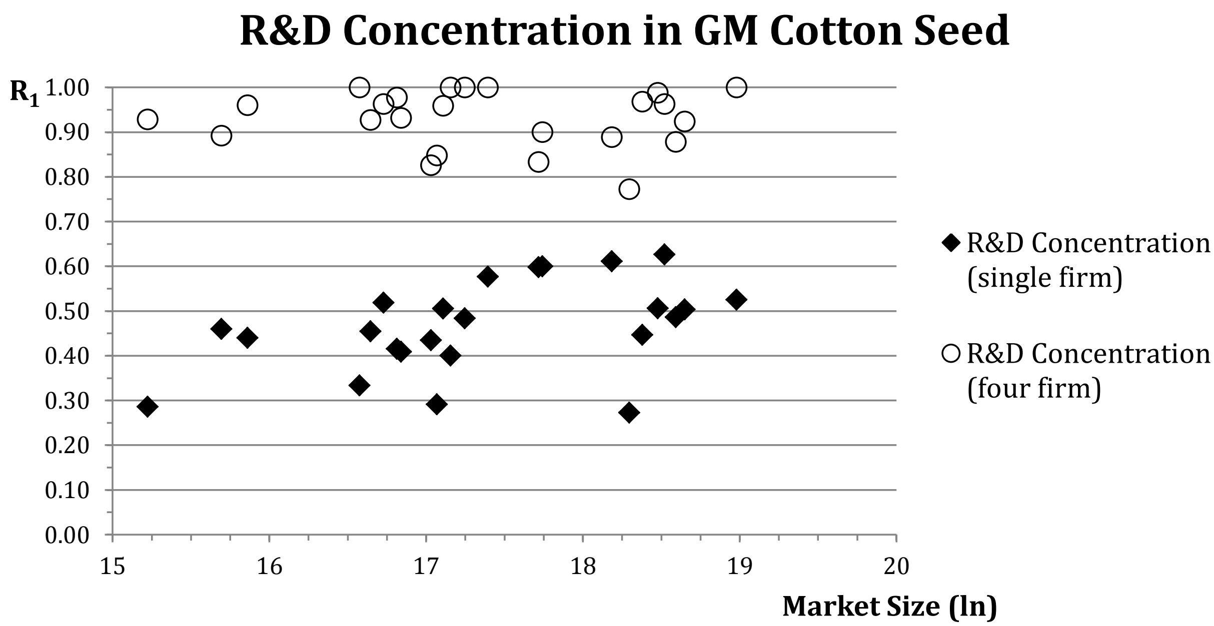


Figure 4: R&D Concentration and Market Size in GM Soybean Seed

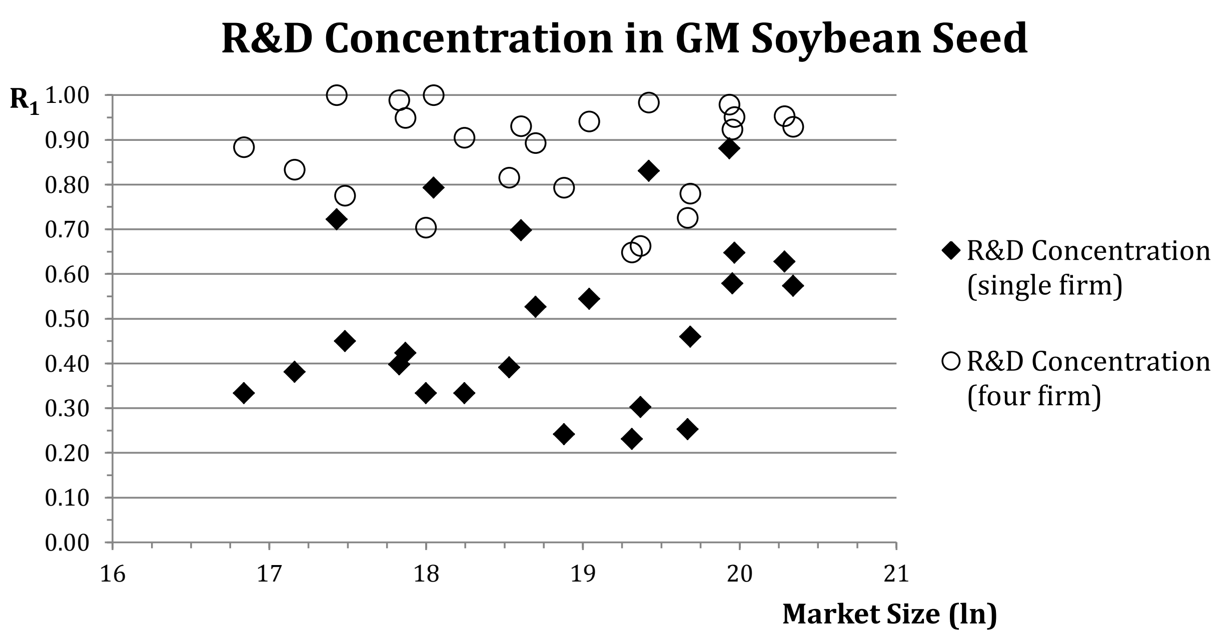


Figure 5: Lower Bound Estimations for R&D Concentration in GM Cotton Seed

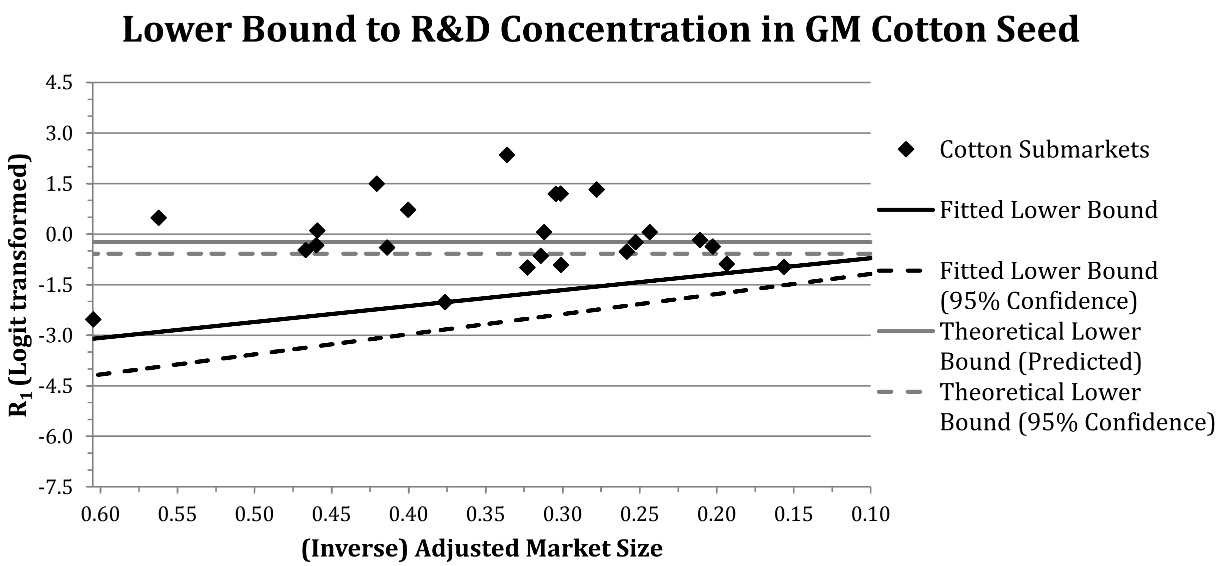


Figure 6: Lower Bound Estimations for R&D Concentration in GM Soybean Seed

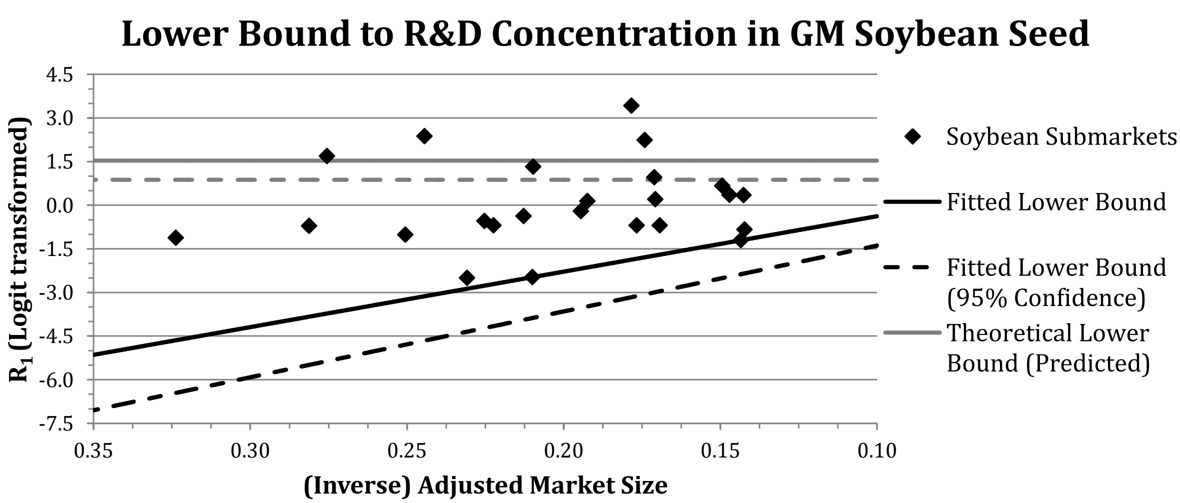


Figure 7: Lower Bound Estimations for R&D Concentration in GM Cotton Seed Adjusted for Mergers and Acquisitions

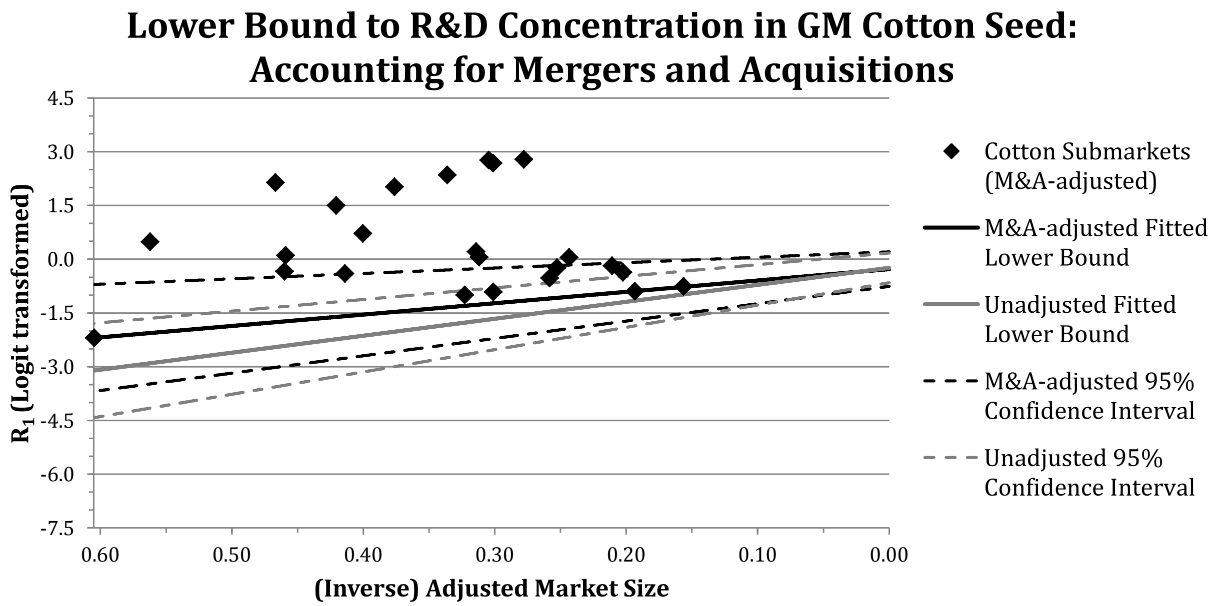


Figure 8: Lower Bound Estimations for R&D Concentration in GM Soybean Seed Adjusted for Mergers and Acquisitions

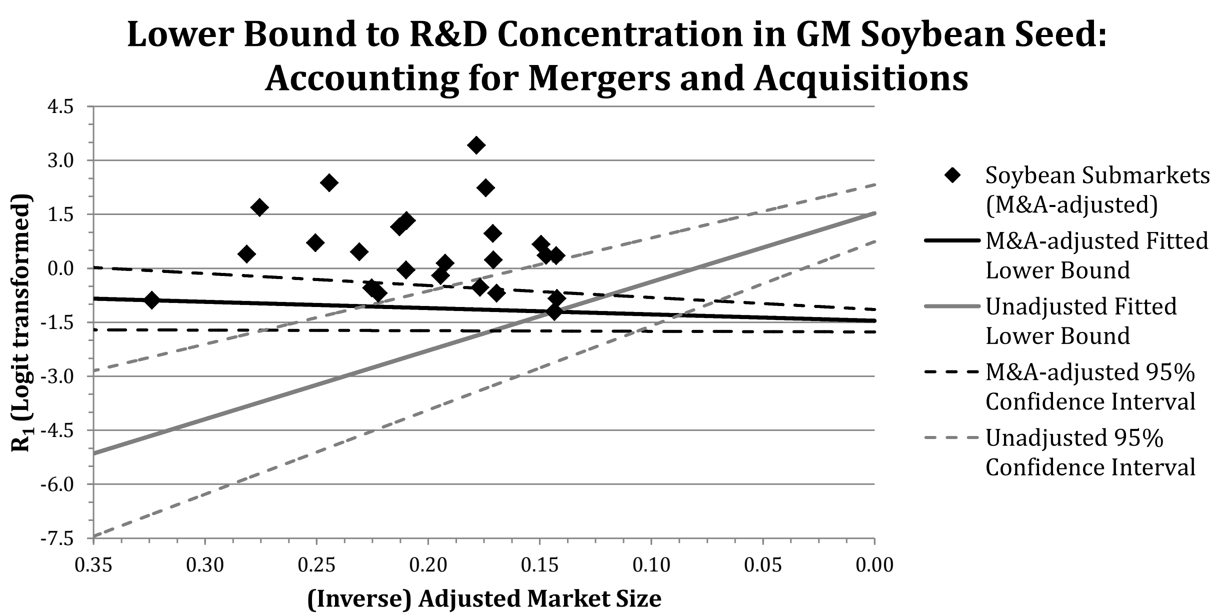


Table 1: GM Seed Submarkets by Crop Type			
	2010 Market Shares (%)	Field Trial Applications	States
Cotton Seed Markets			
Texas	52.42	365	TX
Southeastern States	20.95	611	AL, FL, GA, SC, TN
Mississippi Delta States	10.82	666	AR, LA, MS
Atlantic States	5.87	178	NC, VA
Southern Plains States	5.04	76	KS, MO, OK
Southwestern States	3.19	375	AZ, CA, NM
Classification #3			
Western "Core" States	31.53	1,043	IA, MN, MO, WI
Eastern "Core" States	28.91	1,218	IL, IN, KY, MI, OH
Northern Plain States	22.38	512	KS, NE, ND, SD
Southeastern States	9.40	373	AL, FL, GA, NC, SC, TN
Southern Plains / Mississippi Delta States	9.05	520	AR, LA, MS, OK, TX
Mid-Atlantic States	2.72	318	DE, MD, NJ, NY, PA, VA, WV

Source: Authors' estimates from NASS (2010) Acreage Report and Field Trial Applications

Table 2: Lower Bound Estimations for GM Cotton and Soybean Seed				
Adjusted R&D Concentration Ratio (R_1)	**Cotton Seed**		**Soybean Seed**	
<u>First-Stage</u>				
Adjusted Market Size (θ_1)	-4.735	**	-19.075	**
	(0.719)		(2.080)	
Intercept (θ_0)^	-0.241	**	1.531	**
	(0.200)		(0.381)	
<u>Second-Stage</u>				
Shape Parameter (γ)	1.815	**	1.802	**
	(0.034)		(0.064)	
Scale Parameter (δ)	2.142	**	2.872	**
	(0.057)		(0.076)	
Theoretical Lower Bound (R_1^{∞})^^	**0.473**		**0.786**	
Lower Bound (95% confidence)	0.436		0.681	
Feasible Range ($h \in [0,1]$)	0.440-0.500		0.500-0.822	
Fitted Lower Bound (Largest Submarket)	0.359		0.208	
Likelihood Ratio (χ^2=1)	0.001		0.017	
First-stage Observations	24		24	
Second-stage Observations	22		22	

Standard errors in parentheses.

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

^: Null hypothesis (H_0): $\theta_0 \approx -9.210$.
Null hypothesis (H_0): As the market size becomes large, does the lower bound to R&D concentration converge to (approximately) 0 assuming homogeneity (h = 1)?

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

Table 3: Robustness Checks on Lower Bound Estimations								
	Cotton Seed				Soybean Seed			
Adjusted R&D Concentration Ratio (R_1)	**GM Market Size**	**Homogeneity** (h = 1)	**Setup Costs** (↑ F_0)	**Quadratic Market Size**	**GM Market Size**	**Homogeneity** (h = 1)	**Setup Costs** (↑ F_0)	**Quadratic Market Size**
First-Stage								
Adjusted Market Size (θ_1)	-5.604 *	0.446	-3.875 **	12.241 **	-20.330 **	6.825 **	-17.375 **	-34.876 **
	(2.306)	(1.071)	(0.594)	(1.427)	(1.075)	(1.018)	(1.885)	(6.588)
Adjusted Market Size Squared (θ_2)	-	-	-	-64.333 **	-	-	-	113.046 **
				(6.389)				(22.367)
Intercept (θ_0)^	0.835 **	-1.051 **	-0.302 **	-1.322 **	2.999 **	-2.211 **	1.560 **	1.472 **
	(1.036)	(0.227)	(0.196)	(0.171)	(0.167)	(0.169)	(0.379)	(1.195)
Second-Stage								
Shape Parameter (γ)	1.691 **	2.620 **	1.843 **	1.573 **	1.221 **	1.282 **	1.823 **	1.389 **
	(0.043)	(0.022)	(0.034)	(0.031)	(0.060)	(0.048)	(0.065)	(0.053)
Scale Parameter (δ)	1.913 **	0.996 **	2.121 **	1.774 **	1.584 **	1.295 **	2.986 **	2.164 **
	(0.074)	(0.018)	(0.055)	(0.056)	(0.086)	(0.048)	(0.078)	(0.078)
Heterogeneity Index (h)	0.697	1.000	0.665	0.665	1.000	1.000	0.922	0.922
Theoretical Lower Bound (R_1^{∞})^^	**0.600**	**0.259**	**0.467**	**0.358**	**0.953**	**0.099**	**0.790**	**0.778**
Lower Bound (95% confidence)	0.389	0.192	0.431	0.329	0.938	0.076	0.686	0.386
Feasible Range ($h \in [0,1]$)	0.500-0.697	-	0.425-0.500	0.210-0.500	0.500-9.530	-	0.500-0.826	0.500-0.813
Likelihood Ratio (χ^2=1)	-0.034	-0.226	0.000	-0.035	-0.072	0.018	0.016	0.015
First-stage Observations	18	24	24	24	18	24	24	24
Second-stage Observations	16	22	22	21	16	22	22	21

Standard errors in parentheses.
**,*: Significance at the 99% and 95% levels, respectively.
^: Null hypothesis (H_0): $\theta_0 \approx -9.210$. As the market size becomes large, does the lower bound to R&D concentration converge to (approximately) 0 assuming homogeneity (h = 1)?
^^: Bounds calculated using product heterogeneity for largest submarket for each seed type and infinitely-sized market.

Table 4: Lower Bound Estimations for GM Cotton and Soybean Seeds (Mergers and Acquisitions Adjusted)		
Adjusted R&D Concentration Ratio (R_1)	**Cotton Seed**	**Soybean Seed**
First-Stage		
Adjusted Market Size (θ_1)	-3.173 **	1.752 *
	(0.806)	(0.763)
Intercept (θ_0)^	-0.276 **	-1.456 **
	(0.232)	(0.150)
Second-Stage		
Shape Parameter (γ)	1.232 **	1.664 **
	(0.032)	(0.058)
Scale Parameter (δ)	2.050 **	1.902 **
	(0.079)	(0.055)
Theoretical Lower Bound (R_1^{∞})^^	**0.470**	**0.225**
Lower Bound (95% confidence)	0.427	0.189
Feasible Range ($h \in [0,1]$)	0.431-0.500	.189-0.500
Fitted Lower Bound (Largest Submarket)	0.405	0.262
Likelihood Ratio (χ^2=1)	-0.008	0.013
First-stage Observations	24	24
Second-stage Observations	22	22

Standard errors in parentheses.

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

^: Null hypothesis (H_0): $\theta_0 \approx -9.210$. As the market size becomes large, does the lower bound to R&D concentration converge to (approximately) 0 assuming homogeneity (h = 1)?

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

Table 5: Impact of Mergers and Acquisitions upon R&D Concentration				
	Cotton Seed		Soybean Seed	
Parameter Equivalence				
$t$test: $\hat{\theta}_{0_U} = \hat{\theta}_{0_{M\&A}}$	-0.030		7.429**	
$t$test: $\hat{\theta}_{1_U} = \hat{\theta}_{1_{M\&A}}$	0.386		-9.318**	
$df(\hat{\theta}_0)$	46		30	
$df(\hat{\theta}_1)$	46		29	
Variance Equivalence				
Ftest: $s_U^2(\hat{\theta}_0) = s_{M\&A}^2(\hat{\theta}_0)$	1.896		6.110**	
Ftest: $s_U^2(\hat{\theta}_1) = s_{M\&A}^2(\hat{\theta}_1)$	1.688		7.113**	
	Unadj.	**M&A**	**Unadj.**	**M&A**
Fitted Values				
$\hat{\theta}_0$	-0.241	-0.276	1.531	-1.456
$\hat{\theta}_1$	-4.735	-3.173	-19.075	1.752
Standard Deviation				
$s(\hat{\theta}_0)$	0.949	1.306	1.787	0.723
$s(\hat{\theta}_1)$	3.418	4.440	10.036	3.763
Sample Variance				
$s^2(\hat{\theta}_0)$	0.900	1.706	3.195	0.523
$s^2(\hat{\theta}_1)$	11.682	19.715	100.730	14.161
Observations (N)	24	24	24	24

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

Appendix: Submarket Cluster Analysis

The observable, state-level characteristics that we utilize in the cluster analysis, summarized in Table A.1, can be broadly classified into either data that does or does not vary with crop type and largely is derived from the period prior to the widespread adoption of GM crop varieties. The former includes the state's geometric center for latitude and longitude, climate data including 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 reported by the Current Research Information System (CRIS) for the fiscal years 1990-1995. The data that varies at the state and crop-level includes farm characteristics such as acres planted, number of farms, average farm size, total sales, and average sales per farm reported by the USDA in the 1987 and 1992 US Census of Agriculture prior to the widespread adoption of GM varieties. Additional data on agricultural chemical usage, particularly for soybean farming, were collected from the *Agricultural Chemical Usage: Field Crop Summary* for the years 1990-1995. In order to address dimensionality problems in the cluster analysis, we use principal-component factor analysis to create indices for the temperature and climate data, for the market size data, and for the farm size data. Results of the cluster analysis for cotton and soybean markets as well as robustness checks reveal that both markets can be reasonably divided into six submarkets are summarized in Tables A.2 and A.3, respectively. Additional data from the factor and cluster analyses are presented graphically in Figures A.1–A.15.

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

*: 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

**: 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

***: Cotton - Only AZ, AR, CA, LA, MS, TX; Soybean - Only AR, GA, IL, IN, KY, LA, MS, MO, NE, NC, OH, SD

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

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

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

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

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

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

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

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

Sub-market Analysis: Shares of US Acres Planted

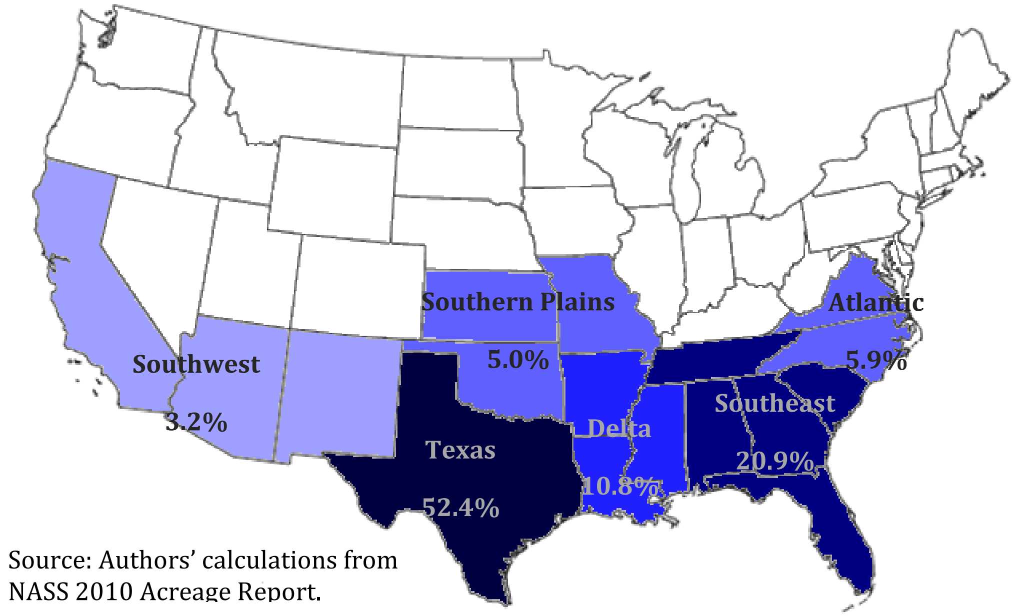


Figure A.1: 2010 Submarket Shares of US Cotton Acres Planted

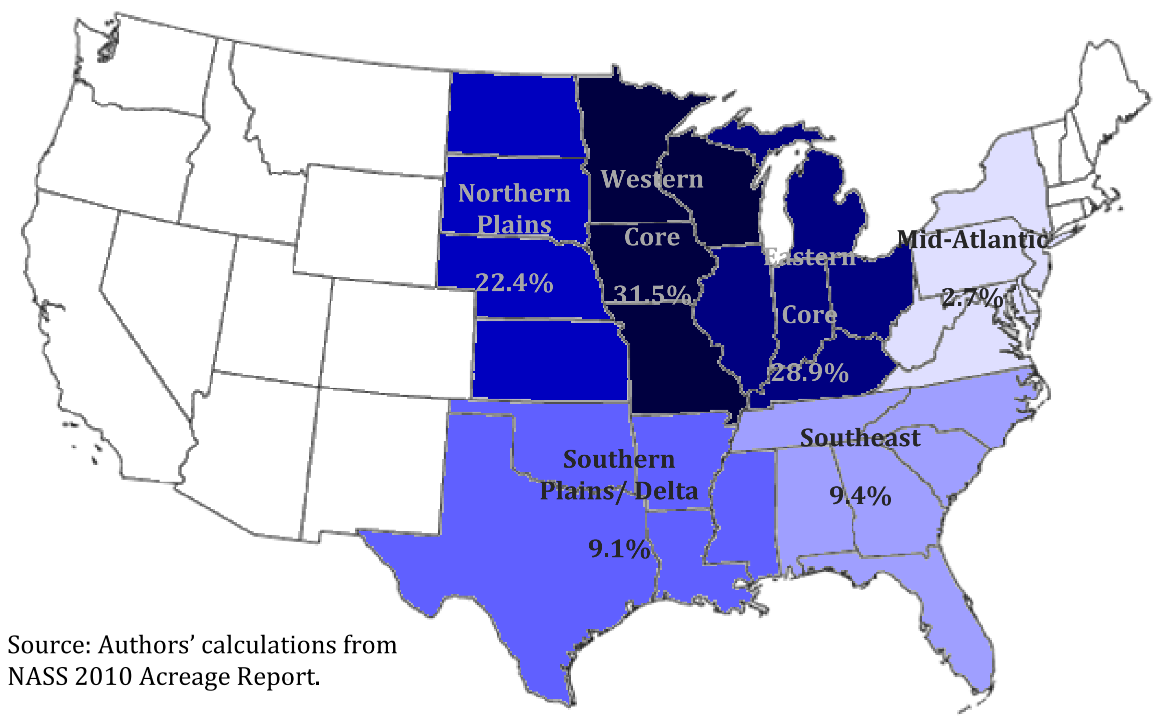
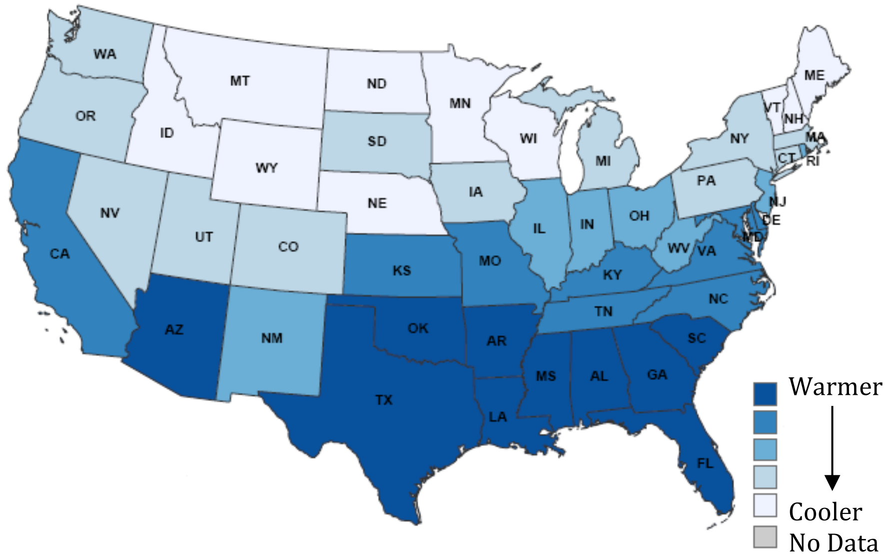


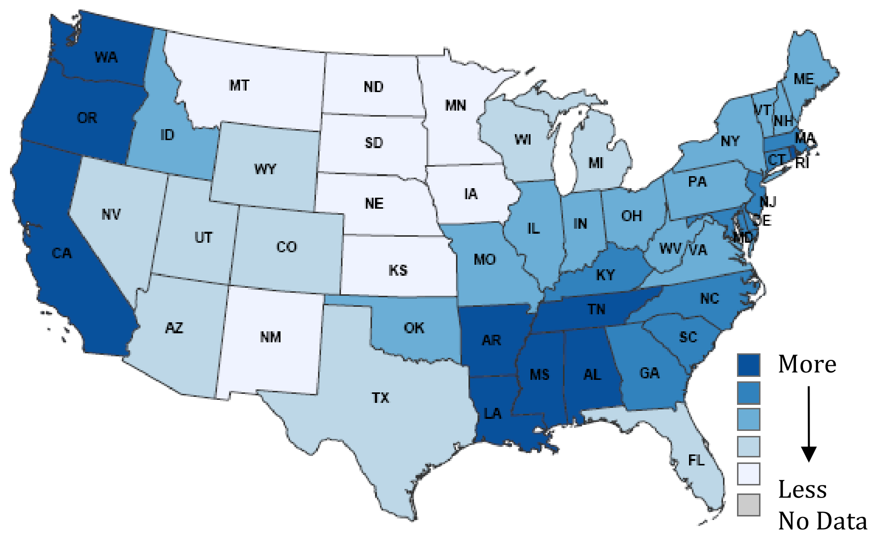
Figure A.2: 2010 Submarket Shares of US Soybean Acres Planted

Sub-market Analysis: State-Level Climate



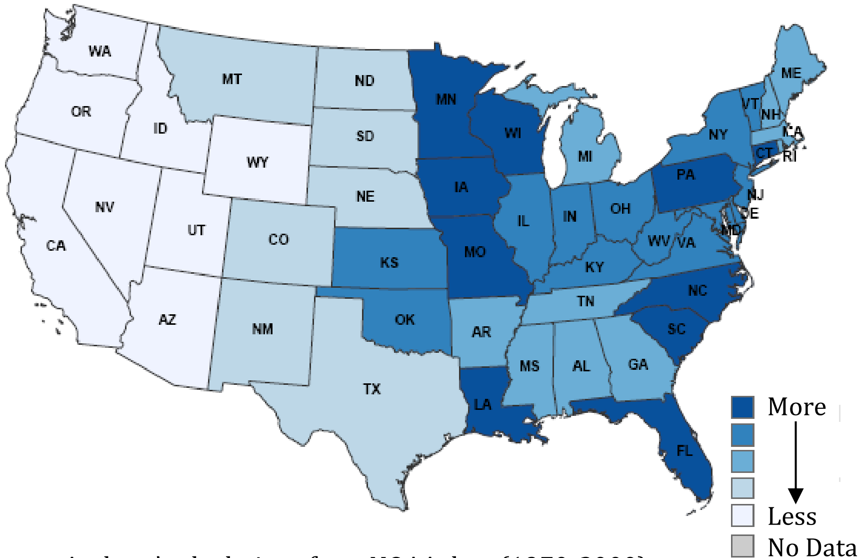
Source: Authors' calculations from NOAA data (1970-2000).

Figure A.3: Average Monthly Temperatures Factor Analysis



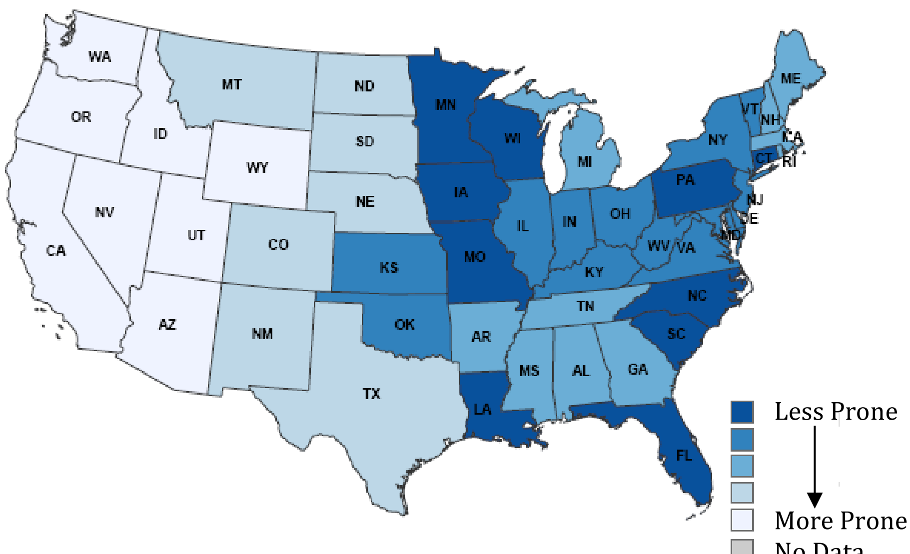
Source: Authors' calculations from NOAA data (1970-2000).

Figure A.4: Average Monthly Precipitation Factor Analysis (1)



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

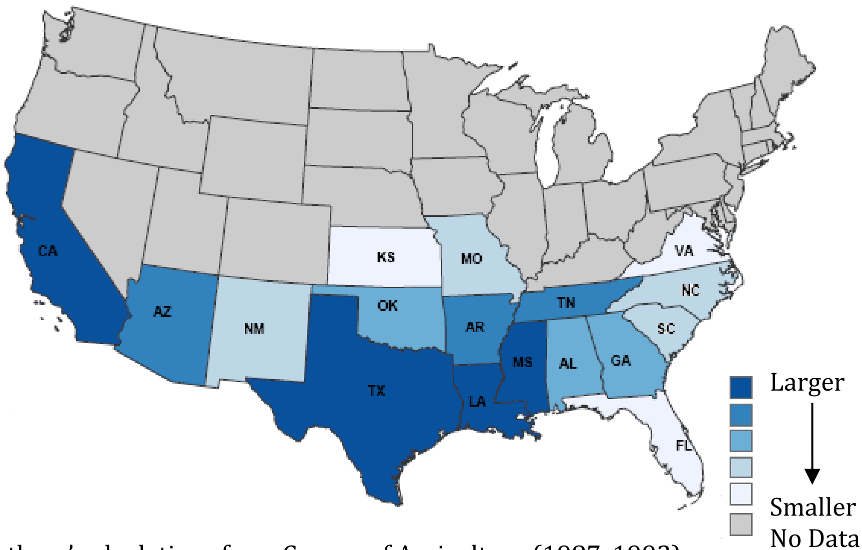
Figure A.5: Average Monthly Precipitation Factor Analysis (2)



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

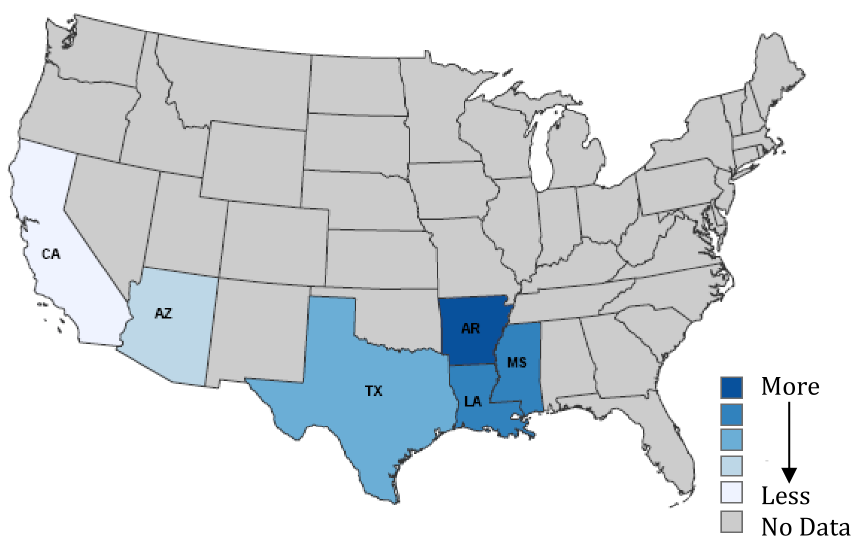
Figure A.6: Average Monthly Drought Likelihood Factor Analysis

Sub-market Analysis: Cotton



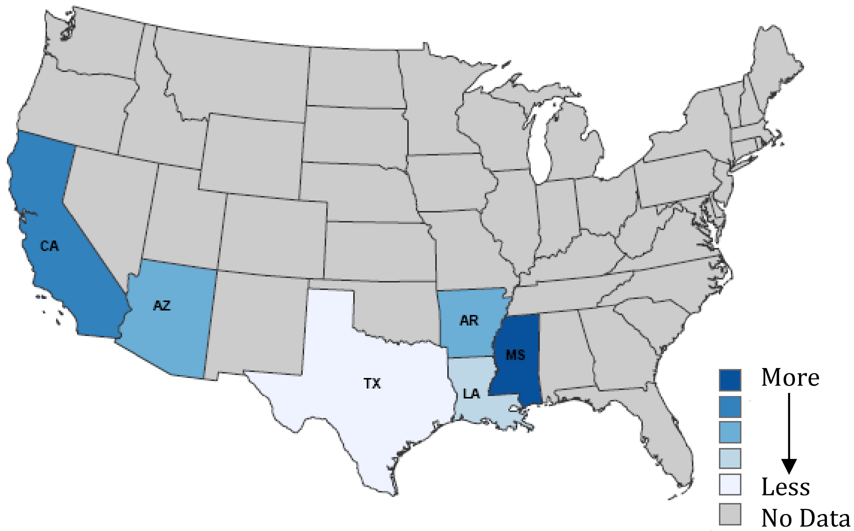
Source: Authors' calculations from Census of Agriculture (1987, 1992).

Figure A.7: Cotton Seed Market Size Factor Analysis



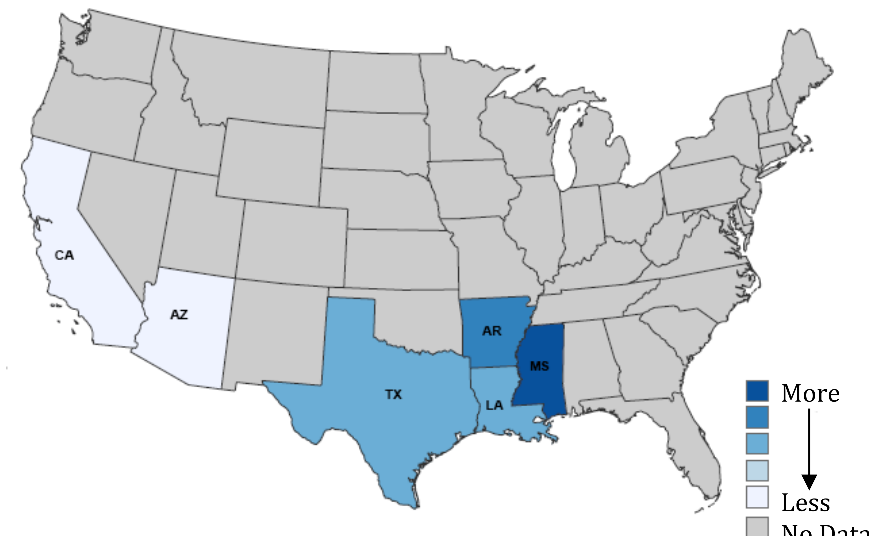
Source: Authors' calculations from Agricultural Chemical Usage (1990-1995).

Figure A.8: Percentage of Planted Cotton Acres Treated with Fertilizer (1)



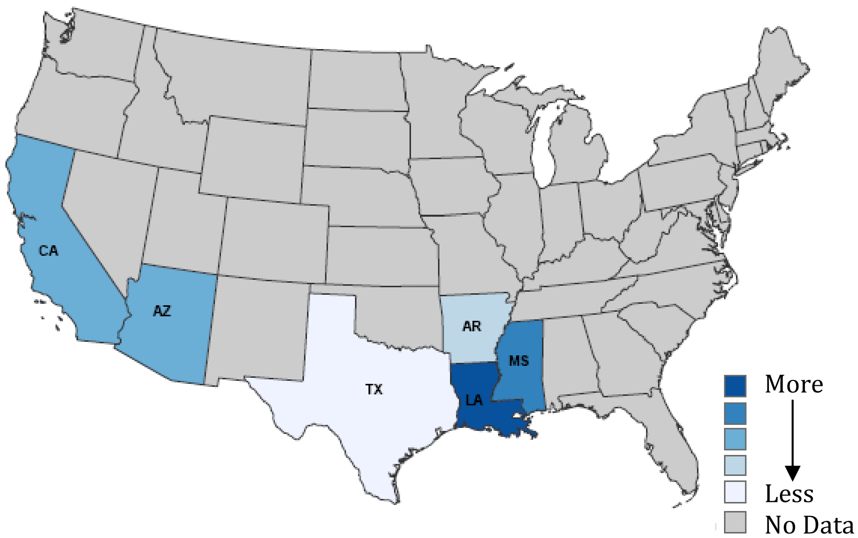
Source: Authors' calculations from Agricultural Chemical Usage (1990-1995).

Figure A.9: Percentage of Planted Cotton Acres Treated with Fertilizer (2)



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

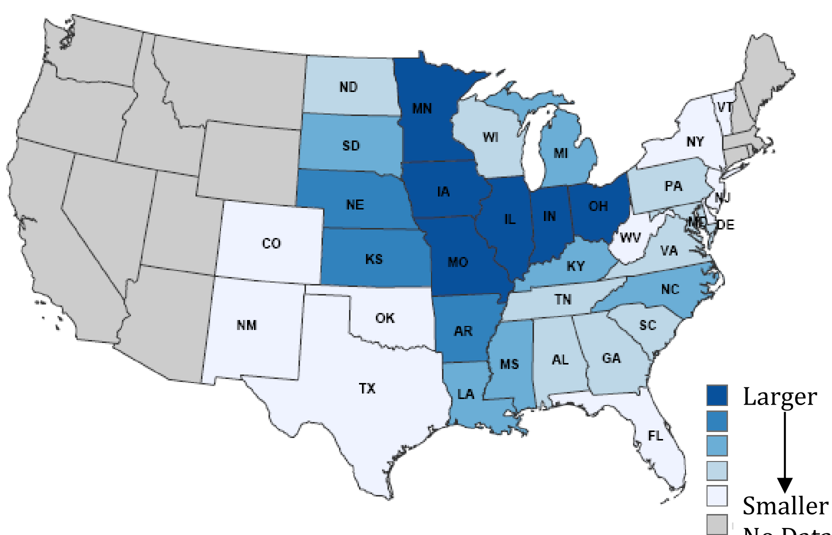
Figure A.10: Percentage of Planted Cotton Acres Treated with Herbicide



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

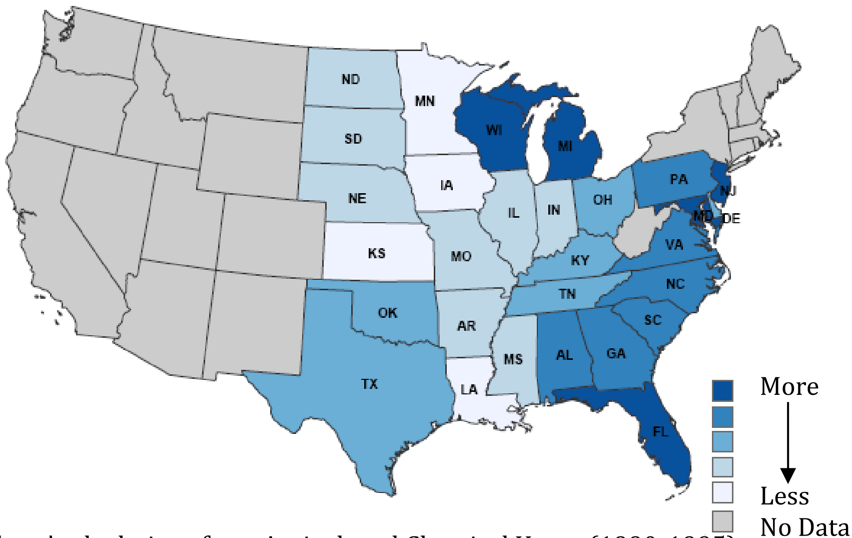
Figure A.11: Percentage of Planted Cotton Acres Treated with Insecticide

Sub-market Analysis: Soybean



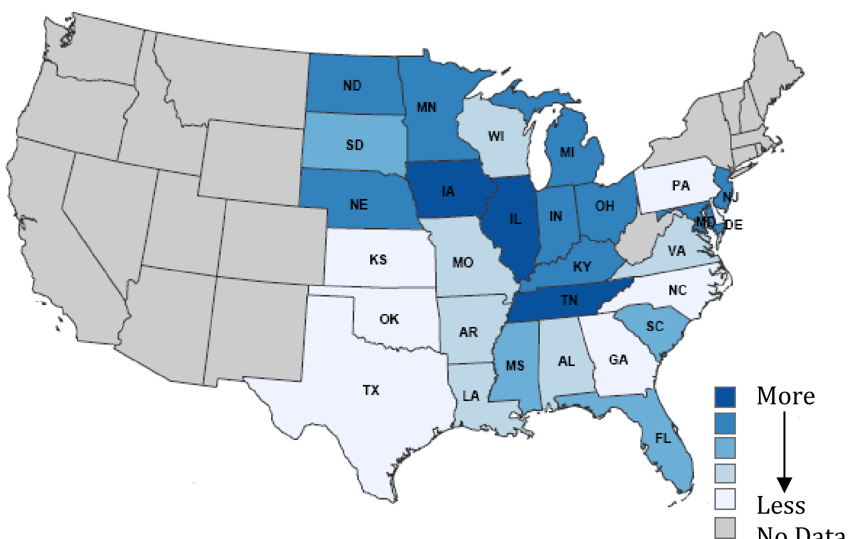
Source: Authors' calculations from Census of Agriculture (1987, 1992).

Figure A.12: Soybean Seed Market Size Factor Analysis



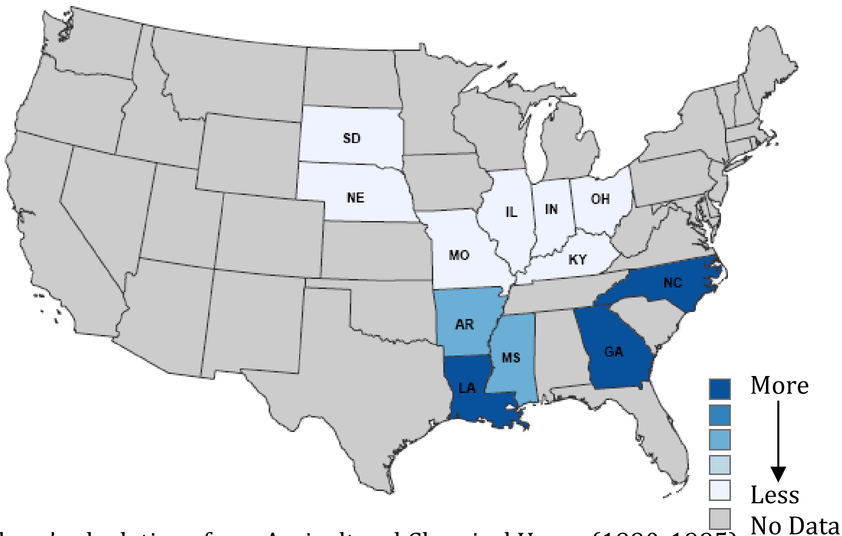
Source: Authors' calculations from Agricultural Chemical Usage (1990-1995).

Figure A.13: Percentage of Planted Soybean Acres Treated with Fertilizer



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

Figure A.14: Percentage of Planted Soybean Acres Treated with Herbicide



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

Figure A.15: Percentage of Planted Soybean Acres Treated with Insecticide

Data Sources

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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