Agribusiness and Applied Economics Report No. 790 **May 2019**

THE INFLUENCE OF STATE-LEVEL PRODUCTION OUTCOMES UPON U.S. NATIONAL CORN AND SOYBEAN PRODUCTION: A NOVEL APPLICATION OF CORRELATED COMPONENT REGRESSION

David W. Bullock, Ph.D.

Department of Agribusiness and Applied Economics
Agricultural Experiment Station
North Dakota State University
Fargo, ND 58108-6050

Acknowledgements

The author would like to thank Edie Nelson who assembled the publication, and William W. Wilson and Bryon Parman who reviewed the publication.

NDSU does not discriminate in its programs and activities on the basis of age, color, gender expression/identity, genetic information, marital status, national origin, participation in lawful off-campus activity, physical or mental disability, pregnancy, public assistance status, race, religion, sex, sexual orientation, spousal relationship to current employee, or veteran status, as applicable. Direct inquiries to Vice Provost for Title IX/ADA Coordinator, Old Main 201, NDSU Main Campus, 701-231-7708, ndsu.eoaa.ndsu.edu. This publication will be made available in alternative formats for people with disabilities upon request, 701-231-7881.

NDSU is an equal opportunity institution.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Over the past 20 years, U.S. agriculture has witnessed profound changes with respect to technology, climate, farm policy, and other factors (ethanol production, Chinese demand, etc.) that have major repercussions with regards to the geographic distribution of crop production — particularly, from a market share and geographic basis. There have been many recent studies that have examined both the direct and indirect impacts of these production factors upon crop yields, acreage, and production from both a temporal and spatial perspective. However, little to no attention has been paid to the impact of these factors upon the relative influence of each individual state's crop production outcomes as they relate to the national outcome.

The purpose of this study is to address this question of state-level geographic importance for U.S. corn and soybeans by employing the following procedure. First, a metric is constructed to measure crop production outcomes at any geographic level by comparing the current year's production to the recent historical norm. This metric, called a production performance index (PPI), is simply the difference between the current year's crop production and the Olympic average (drop minimum and maximum and take arithmetic average of remaining values) of the previous five years of production. The dataset used in the study includes annual crop production values for the 1970 through the 2017 crop years. The PPI, given its five-year lag, is calculated with values for the U.S., each major producing state, and the "Other States" residual from the 1975 to 2017 crop years for both corn and soybeans. The PPI time series is divided into two distinct sets of time periods as a proxy for the changes mentioned above: (1) the 1975 to 1995 crop years, and (2) the 1996 to 2017 crop years. The 1996 crop year was chosen as the dividing point since it represents a watershed year in U.S. corn and soybean production — the commercialization of the first GMO corn (Bt corn) and soybean (Roundup Ready) varieties.

Each states' relative influence upon the national production performance outcome is determined by regressing the individual states' PPI values upon the national PPI value for corn and soybeans under each time period. The regression analysis is conducted using correlated component regression (CCR) – a relatively new statistical tool for sparse and mulicollinear datasets. The absolute value of the standardized coefficient values from the regression model are used to rank each state with regards to its influence. Each state's percentage share of the sum of the absolute coefficient values was also calculated and used to calculate a Herfindahl-Hirschman Index (HHI) by summing the squared values of the percentage shares. The HHI is used as a measure of the geographic dispersion of production importance for the national aggregate.

Overall, the results showed a shifting geographic dynamic for both corn and soybeans with the emphasis shifting from east to west in general direction. This makes intuitive sense as many of the observed technological and climatic changes over the past several decades point towards corn and soybean varieties that require a shorter growing season, and the increase in the number of frost-free days in many of the states in the northern reaches of the U.S. Corn Belt region. Additionally, the greater utilization of irrigation in crop production has likely contributed to the westward expansion of both corn and soybean production — often at the expense of wheat and cotton production. The slight decline in the HHI for corn indicates that production influence is becoming slightly more diversified from a geographic perspective. For soybeans, the opposite effect has occurred with a slight increase in the HHI pointing towards greater influence from the key producing states of Iowa, Minnesota, and Illinois — likely the result of a shift from corn to soybean acres as all three states lost influence shares in corn production between the two time periods.

THE INFLUENCE OF STATE-LEVEL PRODUCTION OUTCOMES UPON U.S. NATIONAL CORN AND SOYBEAN PRODUCTION: A NOVEL APPLICATION OF CORRELATED COMPONENT REGRESSION

David W. Bullock, Ph.D.[1]

INTRODUCTION

During the growing season for major U.S. crops, much of the crop production information generally comes at the state and lower geographical levels. This holds true even through the harvest of a particular crop. For example, the widely following USDA *Crop Progress* reports provide information on crop progress (planting, development, and harvest percentages) and condition (categorical rating percentages) for the major production states for each major U.S. crop on a weekly basis during the growing season. Additionally, soil moisture (sub- and topsoil strata) ratings are reported for the lower 48 states. Projected planted and harvested areas are also reported on a state level for the major states at multiple times during the growing season. Most major private reporting organizations; such as Informa IEG, the Proexporter Network, and INTL FCStone; also generally report growing season production information on a state-level using USDA's major producing states for each major crop.

However, for most economists and industry analysts involved in forecasting, the relevant production numbers are generally at the aggregate national level which feeds into the projected supply and demand balance tables such as those regularly reported in the USDA World Agricultural Supply and Demand Estimates (WASDE) which is reported monthly throughout the year. This raises the natural question of whether certain states production outcomes provide a bellwether signal for the national production outcome. That is, if a particular state (or group of states) is experiencing excellent (above average) growing conditions, does this translate into a national production outcome that is above average?

Besides the aforementioned economists and industry analysts, this question also has importance for many firms engaged in agricultural production and agribusiness. For example, firms involved in logistics are interested in this information as they have to plan ahead in terms of positioning transportation resources to meet the anticipated demands for transportation from the points of production to the sources of demand. Additionally, this information is also important from an inter-temporal perspective as storage decisions can effectively utilize this type of information.

Additionally, firms involved in agricultural production technology can use this information in making long-term planning decisions. This includes the orientation of investments in technology development at major public research institutions such as land-grant universities, in addition to the rapid increase in private sector spending on research and development.

Finally, this question has very important implications regarding risk management. In particular, have changes in the crop production technology, climate, farm policy, and other production

[1] Research Associate Professor, Department of Agribusiness and Applied Economics, North Dakota State University. Address all correpondence to: David W. Bullock, 624 Barry Hall, 811 2nd Avenue North, Fargo, ND 58102 or david.w.bullock@ndsu.edu.

factors over the past twenty years resulted in greater or lesser production risk from a geographical perspective. In particular, many of the risks impacting national crop production (droughts, disease outbreaks, etc.) tend to be regional rather than national in nature. If a particular region has a highly significant influence upon the national outcome, then overall production risk from these events is likely to be increased rather than reduced. Diversification is an important risk management strategy and the more diverse the geographic importance of crop production, the lower the risk posed to the national aggregate.

It is a well known fact that changes in crop production technology, farm policy, and global climate are having profound effects upon not only the level but the geographic distribution of crop production in the U.S. and around the world. For example, since the first half of the 1990's, the acreage planted to soybeans in North Dakota has increased by over 800 percent – an increase of over 5.2 million acres which moves North Dakota from #18 to #9 in production ranking over the same period.

A recent article in the *Wall Street Journal* (Bunge 2018) drives home the effect of climate change and shorter maturing crop varieties by examining the changes in crop production occurring in Upper Alberta, Canada. Since 1950, average temperatures around La Crete, Alberta have increased by 3.6 degrees Fahrenheit which has increased the growing season by nearly two weeks. While wheat and canola still dominate crop production in Canada, the area planted to corn has increased by 20 percent and soybeans has roughly doubled over the past decade alone. The article quotes Cargill CEO David MacLennan from a 2016 interview:

> *"Today, the U.S. corn belt is in Iowa, Illinois, Indiana. In 50 years, it may be in Hudson Bay, Canada."*

The main purpose of this study is to devise a methodology for measuring the relative importance of state-level production outcomes in predicting the national U.S. outcome for two crops: corn and soybeans. First, a production performance metric is constructed that measures the annual production level versus the normally observed level over the previous five years. This metric is constructed for each of the major corn and soybean production states (18 each) as determined by USDA-NASS in their monthly and annual *Crop Production* reports. This metric is also calculated for the total production of the United States as a whole and the residual amount (i.e., United States minus 18 states total) is used to calculate the performance for the states ("Other States") not included in the set of 18 major producing states. These metrics are calculated for the 1975 through 2017 crop years using historical data (1970 to 2017) from the USDA-NASS.

To measure the relative influence of the individual states upon the national crop production outcomes, regressions were set up for each crop and time period (4 in all) with the United States production performance metric as the dependent variable and the individual states' production performance metrics as the explanatory variables. The regression standardized coefficient values from each regression are used to rank each state in terms of its influence upon the national metric. Because of data sparsity and multicollinearity issues, the standard OLS regression model could not be used to derive accurate coefficient estimates. Therefore, a relatively new regression procedure, called *correlated coefficient regression* (CCR; Magidson 2010), that was developed to directly handle the issues of data sparsity and multicollinearity, was applied to the dataset.

A secondary purpose of this study is to measure the impact of recent technological, climatic, and policy factors upon the state ranking and also the geographic concentration of this influence. To accomplish this purpose, the dataset containing the performance metrics is divided into two time periods: (1) the 1975 to 1995 crop years representing the period just prior to the commercialization of the first GMO crop varieties, and (2) the 1996 to 2017 crop years which covers the period following the commercialization of GMO varieties. The latter period can also be characterized by the increasing influence of bio-fuel production, major changes in farm policy moving away from supply control to a more market-based income support emphasis, and the highly publicized increases in global temperatures and weather volatility which had been occurring even prior to 1996.

To measure the geographic concentration of influence, a Herfindahl-Hirschman Index (HHI) was calculated on the absolute percentage shares of the standardized coefficients. By comparing the derived HHI from each time period for each crop, the impacts of the technological, climatic, and policy factors upon the concentration in production risk can be observed.

In the next section, background on the recent developments in the aforementioned production factors is provided. This is followed by a review of the literature pertaining to the impacts of these factors upon crop yields and the geographic distribution of acreage and production. The next section describes the data and the research methodology utilized in this study, followed by the research results, and finally a section discussing the major findings and conclusions.

BACKGROUND

The 1996 crop year represented a watershed in the history of U.S. corn and soybean production. The prior year (1995) saw the introduction of the first approved GMO crop with the introduction of the Flavr Savr tomato by Calgene, Inc. (Bruening and Lyons 2000). The 1996 crop year saw the commercialization of the first GMO variety offering herbicide tolerance (HT) which was introduced by Monsanto (Roundup Ready™ soybeans). The same year saw the commercialization of the first variety offering insect resistance (IR) was introduced by Ciba-Geigy (Maximizer™ corn) which contains a gene that expresses a protein from the bacterium Bacillus thuringiensis (Bt) which is toxic to the European corn borer and related species (James and Krattiger 1996).

Until 2009, herbicide tolerance and insect resistance remained the dominant traits in GM development (Shakya, Wilson and Dahl 2013). These traits were often offered in stacks (multiple GM traits in one variety) and were primarily producer focused in terms of their benefits. Additional producer traits under development include drought tolerance (DT) and nitrogen-use efficiency. While having the potential to boost yield, most of these early traits focused primarily upon reducing production risk and producer costs as well as providing convenience.

According to data from USDA-ERS[2], approximately 92% of U.S. corn acreage and 94% of U.S. soybean acreage was planted to GE crops in 2017. This is up from 25% of U.S. corn acreage and 54% of U.S. soybean acreage in 2000. All of the soybean acreage was planted to HT varieties in 2017 while 84% of the corn GE acres was planted to stacked trait varieties, 13% to HT only varieties, and the remaining 3% to IR only (Bt) varieties. This differs considerably from 2000

[2] https://www.ers.usda.gov/data-products/adoption-of-genetically-engineered-crops-in-the-us/

when 72% of corn GE planted acres was IR only, 24% was HT only, and 4% was stacked varieties for corn. Figure 1 shows the planted acreage share history for U.S. GE corn and soybean acreage for the 2000 to 2017 crop years.

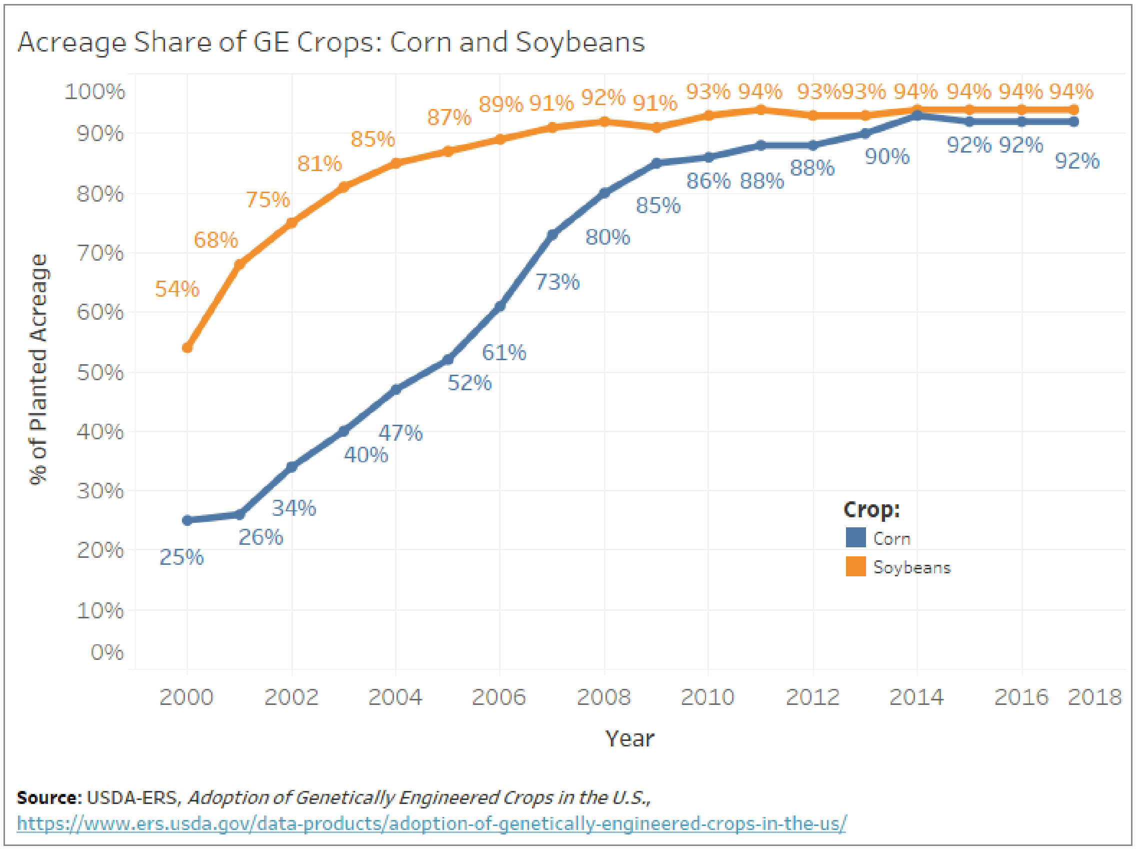

Figure 1. Acreage Share of Genetically Engineered Crops: U.S. Corn and Soybeans

In addition to the aforementioned GMO technological advances, other advances in varietal development and crop breeding have had a major influence upon the level of crop production and its geographic distribution. For example, the development of faster maturing corn and soybean varieties has drastically shortened the required growing season and increased the reach of corn and soybean production into northern states with shorter growing seasons. Other areas of technological advancement in corn and soybean production include precision agriculture, improved tillage practices, and fertilization.

Climate change has also had a major impact upon the level and geographic distribution of U.S. corn and soybean production. The changes in the earth's climate in the past century have been well-documented. According to data compiled by NASA[3], the current level of CO2 in the atmosphere is nearly 100 million parts per million higher than the highest level from the previous 400,000 years (using data gleaned from ice core samples). Largely driven by the much higher CO2 levels, the average planetary surface temperature has risen 2 degrees Fahrenheit since the late 19th century with 16 of the 17 warmest years on record since 2001. Additional data

[3] https://climate.nasa.gov/evidence

documents warming ocean temperatures, decreasing snow cover, rises in sea level, declining Arctic sea ice, a notable increase in extreme weather events, and increased ocean acidification. These trends are documented in greater detail in a recent report by the leading international body on climate change (IPCC 2015).

Changes in U.S. farm policies and crop insurance have also contributed to the changing geographic nature of U.S. corn and soybean production. The passage of the Federal Agricultural Improvement and Reform (FAIR) Act of 1996 (also informally called the "Freedom to Farm Act") greatly increased producer planting flexibility by allowing participants to plant 100 percent of their total base acreage to any crop with some limitations on fruits and vegetables. This effectively allowed farmers in traditionally non-corn and soybean producing regions (such as the wheat producing areas of the Northern Plains) to experiment with planting short-season corn and soybean varieties. However, the statistical evidence is relatively weak regarding the impact of "decoupled" payments upon U.S. crop area (Adams et al. 2001).

Reforms to the crop insurance program and the introduction of new revenue insurance products due to passage of the Federal Crop Insurance Reform Act of 1994 greatly reduced the production and marketing risks faced by row crop farmers by making the coverage more affordable (via premium subsidies and catastrophic coverage) and reducing the impact of production risk in forward contracting and/or hedging of crops (via revenue insurance). This reduction in overall production risk may have had the side effect of reducing the value of enterprise diversification as a risk management strategy (O'Donoghue, Roberts and Key 2009) resulting in changes in the geographical distribution of crop acreage.

Finally, the emergence of ethanol, due to a combination of rising gasoline prices and Federal policies supporting bioenergy, as a significant demand market for corn has also had an impact upon the distribution of acreage to both corn and soybean production (Wallander, Claassen and Nickerson 2011). Farm-level data indicate a net expansion of both corn and soybean acreage at the expense of cotton and uncultivated hay acreage.

PREVIOUS STUDIES

For major U.S. field crops, total production at all geographic levels can be broken down into two components: (1) the yield per harvested acre, and (2) the number of harvested acres. Harvested acreage is generally a function of planted less abandoned acreage. In the long term, changes in technology, climate, and policy will have an influence upon the geographic distribution of both planted acreage, and yield per planted acre. Much of the previous research in this area can be divided into studies that examine these changes from either the perspective of impacts upon crop yields or the geographic distribution of planted acreage and crop production.

Impacts of Climate, Technology, and Policy upon Crop Yields

Thompson (1963, 1969, 1970, 1986, and 1988) built upon the earlier work of Wallace (1920), Ezekiel (1941), Houseman (1942), and Hendricks and Scholl (1943) in using regression models to separate the impacts of technology and weather upon yield. Thompson used piecewise linear trends to model technological advancements in corn and soybean production. The fitted yield values from these models, referred to as "normal weather yields", were used as inputs into curvilinear regression models that examined the impacts of seasonal precipitation and temperature (measured as deviations from normal) to explain yield performance. One of the major geographic conclusions from Thompson's work was the overlap in the optimal seasonal weather conditions for growing both corn and soybeans – the main exception being the importance of August precipitation to soybeans (important in the pod-filling stage). This led to the prediction, in 1970, that the Corn Belt region would eventually be referred to as the 'Corn and Soybean Belt' region.

Thompson also examined the long-term effects of climate change upon crop yields, noting that the observed trend of increasing CO2 in the earth's atmosphere is likely a contributing factor in increases in corn yields (1986). He examined the impacts of increased weather variability, beginning in 1972, upon corn yields noting that there was also greater variability in yields following 1972. He indicated a cyclical pattern in corn yields of approximately 18 years in length that may be tied to lunar and El Nino cyclical phenomenon; however, no conclusive evidence of a connection was established (1988).

Menz and Pardey (1983) examined the impact of nitrogen fertilization upon corn yields and the question of whether yields were potentially reaching a plateau. They used a regression model that effectively split the impact of nitrogen fertilization rates from weather and other technological impacts. They used trend as a proxy for other technologies, weighted average July precipitation in five major corn belt states as a proxy for weather, and a dummy variable to separate out the effect of the 1970 corn blight. Their results showed a much reduced but still positive marginal physical product (MPP) from nitrogen application (dropping from 0.79 bu/ac/lb in 1954-60 to 0.15 bu/ac/lb in 1971-80). They found a constant contribution of 1.0 bushels per year attributable to non-nitrogen technologies. They also noted the disappearance of a price response by corn yields, finding that the earlier results of Houck and Gallagher (1976) no longer appeared to hold.

A more recent analysis of the long-term impact of technology and climate change upon corn and soybean yields can be found in two studies by Tannura, Irwin and Good (2008a, 2008b). Using a modified version of Thompson's (1988) multiple regression model, they examined three key questions: 1) Has the relationship between weather, climate change, and technology, and corn and soybean yields in the key producing states (Illinois, Iowa, Indiana) changed since the last comprehensive studies? 2) Has the trend rate of yield growth for corn accelerated since the mid-1990's? and 3) How does the accuracy of yield forecasts from the regression model compare to benchmark forecasts such as those generated by the USDA?

In addressing the first question, their results showed an asymmetry in that yields (relative to trend) were more negatively impacted by unfavorable weather versus positive impacts due to favorable weather conditions. Technology, June/July precipitation, and July/August temperatures were all found to be significant in impacting the corn yield. May/June temperatures were found to be minimal in significance. For soybeans, yields were most impacted by technology and June through August precipitation (August being most significant). July/August temperatures were also found to be important in soybean yield performance.

To further address the first question, the authors utilized two statistical tests that were applied to the modified Thompson models to see if structural changes in the relationships had occurred. The first set of tests used the unknown breakpoint tests of Quant (1960) and Andrews (1993) to derive the Quant Likelihood Ratio (QLR) statistic. The tests identified potential breakpoints in 1988 for Illinois corn, 1983 for Iowa corn, and 1988 for Iowa soybeans. To better assess the reasons for these structural changes, a second set of tests using dummy variables was applied to the explanatory variables in the models. These results were inconclusive in identifying an obvious reason for the observed structural changes. Then, the QLR test was applied to groupings of the explanatory variables. The July/August temperature variable for Iowa corn had a significant breakpoint in 1983. However, the authors generally conclude, given some of the unreasonable implications of the breakpoint analysis and the fact that impacts would likely be seen across both crops and all three states, that the structural change tests were inconclusive in terms of a change in the relationship between yield and the explanatory variables.

To address the second question of technology acceleration impacts upon corn yields, the authors deployed two forms of structural change tests to the corn yield trendline model alone. The first set of tests utilized the QLR unknown breakpoint test of structural change. These tests failed to identify the mid-1990's as a breakpoint in the slope of the corn yield trend model. The second set of tests used the Chow test of a specified structural break at a specified point (using 1995 as the breakpoint). These tests also indicated that there was not significant structural break in the corn yield trendline model in the mid-1990's.

To address the third question, the authors set up a yield forecasting competition between the modified Thompson model, a yield trendline model, and the USDA forecasts for key forecasting dates (June 1, July 1, August 1, and September 1) during the growing season for both corn and soybeans from 1980 through 2006. The forecasts were compared using the root mean squared error (RMSE), root mean squared percentage error (RMSPE), the mean average error (MAE), the mean average percentage error (MAPE), and the modified Diebold-Mariano (MDM) test as developed by Harvey, Leybourne and Newbold (1998). The modified Thompson model forecasts were no more accurate than the trendline forecasts for June 1 and July 1. For corn, the modified Thompson model began outperforming the trend model beginning on August 1 while for soybeans, the model outperformed trendline on September 1. This make sense since the results of the modified Thompson model should improve as more weather information is available during the growing season. USDA forecasts outperformed both the modified Thompson and trendline models over all time periods. However, the encompassing tests showed that combining the modified Thompson with the USDA forecast could improve forecast accuracy an average of 10 percent for corn and 6 percent for soybeans.

Kukal and Irmak (2018) examined the long-term variability in climate and yields for corn, soybeans, and sorghum using 46 years (1968 to 2013) of county-level data from 9 Great Plains states (Kansas, North Dakota, South Dakota, Wyoming, Iowa, Nebraska, Oklahoma, Texas, Colorado). Analysis of crop yield variability was limited to only those counties that had data available across at least 60% of the study period (246 for corn, 242 for sorghum, and 225 for soybeans). Yield and climate variability were both represented using the statistical coefficient of variation. Using the R^2 statistic from regressing climate variability (temperature and precipitation) upon yield variability, they found that climate variability explained 18%, 23%, and 23% of yield variability for corn, sorghum, and soybeans over the study period. They found that the general trend in temperatures was beneficial for corn but detrimental for sorghum and soybeans. The general trend in precipitation was beneficial for all three crops. They examined the differences between irrigated and non-irrigated yields and found that irrigated yields were more robust and an effective mitigation strategy against climatic impacts. They also found considerable geographic variation in the results regarding climatic trends and the yield responses to those trends.

There have been numerous other studies examining the impacts of climate change and technology upon yields. Garcia et al. (1987) assessed the impacts of weather and technology upon U.S. corn yield variability over two time periods: 1931 to 1960 and 1961 to 1982. They found that when yields were adjusted for weather, the variance in corn yield is more likely to be equal across the time periods – indicating the relative importance of weather in explaining yield variability.

Kaufmann and Snell (1997) estimated a hybrid model that accounted for both climatic and social (market conditions, technical factors, scale of production, and policy environment) impacts upon corn yields in the U.S. They found that their model was highly effective in assessing the social impacts upon corn yields – particularly in assessing the relative costs and benefits for adaption.

Schlenker and Roberts (2009) examined the nonlinear and asymmetric relationship between temperature and crop yields and examine the implications of long-term temperature projections. They found, holding current growing regions fixed, that area-weighted average yields are predicted to decrease by 30 to 46 percent before the end of the 21st century under the slowest warming scenario and decrease by 63 to 82 percent under the most rapid warming scenario using the Hadley III model.

Lobell, Schlenker and Costa-Roberts (2011) developed a database of yield response models to evaluate the impact of climate trends on global yields by country for the period of 1980 to 2008. The crops analyzed were corn, wheat, rice, and soybeans. To estimate the impacts of climate trends, four scenarios were applied to historical temperature and precipitation data for each country over the study time period: (1) actual values over the time period, (2) actual temperature and detrended precipitation, (3) detrended temperature and actual precipitation, and (4) detrended temperature and precipitation. For corn and wheat, the results indicated a global net yield loss of 3.8 and 5.5 percent respectively over the study time period. For rice and soybeans, the results essentially balanced out across regions.

Cai et al. (2013) developed a climate index using principal components to estimate the linkage between climate and regional crop yields in the U.S. The indices were also forecast using three long-term climate projections (Australian CSIRO 3.5, Canadian CGCM 3.1, and Japanese MIROC 3.2) to derive long-term implications for crop yields. Their results indicated that future hotter/drier weather conditions will likely have a more significant negative impact upon crop yields in the southern states with only mild impacts in the most northern states.

Lobell et al. (2014) examined field level data on corn and soybean yields in the Central U.S. from 1995 to 2012 to examine changes in yield sensitivity to drought. Their results indicate that yield sensitivity to drought has been increasing over time. The authors suggest that a key factor contributing to this increased sensitivity may be planting density even though new varieties are more robust to crowding.

Tolhurst and Ker (2015) modeled the impact of technological change upon crop yields using a mixture of normal distributions (EM algorithm) with embedded trend functions to account for technology in the different components of the distribution. This allows for the incorporation of the higher moments (beyond first two moments) of the yield distribution in the estimation of the impacts of technological change. The model was applied to county-level corn, soybean, and wheat yields from Illinois, Indiana, and Iowa for the period from 1955 to 2011. Their results indicated that technological change may alter the shape of the yield distribution beyond just a location-scale shift in the parameters. In particular, they found that the rate of technological change in the upper components was greater than the rate of change in the lower components of the yield distribution for all three crops.

Du et al. (2015) examined the impact of exogenous geographic and climatic factors upon the moments of the crop yield distribution with a focus on the skewness (3rd moment) parameter. The data examined included crop insurance unit yield data for corn and soybeans in 13 states (Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, Oklahoma, South Dakota, and Wisconsin) and wheat yield data from 11 states (all mentioned except Iowa and Wisconsin) from 1990 to 2009. Their results indicate that better natural resource endowments (climate and soils) decrease the observed skewness in yields which supports their theoretical hypotheses.

Chen, Chen and Xu (2016) examined the impact of climate change upon corn and soybean yields in China using the Schlenker and Roberts (2009) model. Their results indicated a loss of about $820 million to China's corn and soybean sectors in the previous decade due to global warming. They also projected that China's corn and soybean yields will decline by 3-12% and 7-19% respectively by 2100 if current climate trends continue.

Huffman, Jin and Xu (2018) examined over a half-century of panel data on U.S. Midwest rain-fed state-average corn yields. They broke down the observed yields into two components: (1) the yield potential which was modeled as a stochastic production frontier where nitrogen fertilization, public corn research, and biotechnology introduction have an effect; and (2) damage to yield potential due to weather and pests which was modeled as an asymmetric control function. They found that nitrogen use, public corn research, and adoption of biotechnology increase yield potential while soil moisture stress reduces yield potential. They also found that excess heat severely reduces nitrogen productivity and that biotechnology primarily abates yield damage due to soil moisture stress but does not abate damage due to excessive heat.

Impacts of Climate, Technology, and Policy Upon The Distribution of Crop Acreage and Production

Much of the research focusing upon the impact of climate and technology upon the geographic distribution of crop acreage and production has generally linked these impacts with other direct or indirect economic impacts. Rosenzweig and Parry (1994) combined data from individual crop yield studies to obtain a global picture of the simulated change in crop yield associated with different climate scenarios. A world food trade model was then applied to simulate the economic consequences of the potential changes in crop yields and global production levels. A major conclusion of their study is the appearance of a major disparity in agricultural vulnerability between developed and developing countries with regards to climate change.

Mendelsohn, Nordhaus and Shaw (1994) measured the economic impact of climate change upon land prices using a Ricardian approach rather than a production function approach. Using cross-sectional data from almost 3,000 counties in the U.S., they find that higher seasonal temperatures (with the exception of autumn) result in lower farm values while increasing seasonal precipitation (outside of autumn) increases average farm values. Applying the Ricardian approach did result in a significantly lower estimated impact when compared to the production function approach with one case even suggesting (excluding positive impacts from higher CO2 levels) that global warming may have significant economic benefits for agriculture. They also found that the most negative impacts of climate change upon cropland values would occur in the southern United States with most of the positive impacts occurring in the central and northern states.

Schlenker, Hanemann and Fisher (2006) utilized a hedonic (Ricardian) regression approach to examining the link between climatic variables (degree days and precipitation) and U.S. farmland values east of the 100th meridian (the boundary of agriculture not primarily dependent upon irrigation). They utilized a hedonic model to project farmland values under four long-term climatic scenarios from the Hadley HadCM3 model (B1, B2, A2, and A1F1). Their results generally show both positive and negative impacts upon farmland values on a county-by-county basis. The number of counties exhibiting losses exceeded the number of counties exhibiting gains.

Deschenes and Greenstone (2007) estimate the impact of climate change upon the U.S. agricultural sector by estimating the impact of year-to-year variations in temperature and precipitation upon agricultural profits. They used county-level panel data to estimate the effect of weather on agricultural profits, conditional on county and state by year fixed effects. Using the long-run climate change predictions from the Hadley 2 Model, they found that climate change will result in a $1.3 billion (2002 $) or 4.0 percent increase in aggregate U.S. agricultural profits. Their results show significant variation across states in terms of changes in profitability with South Dakota (+$720 million) showing the largest projected increase and California (-$750 million) showing the largest decrease. In percentage terms, West Virginia (+189.6%) had the largest increase and New Hampshire (-127.4) had the largest decrease.

Marshall et al. (2015) examined the impact of climate change on the regional water balances in the U.S. and the resultant impact upon the geographic distribution of irrigated versus non-irrigated acreage. They explored farmer response under two hypothetical cases: (1) where irrigation water supply constraints don't exist so that production impacts are purely due to biophysical factors, and (2) where decisions are constrained by irrigation water supplies to isolate the impacts of water balances. Their results indicated that, from a national perspective, the impact of irrigation water supply distribution was small relative to the direct biophysical impacts upon crop yields.

Miao, Khanna and Huang (2016) investigated the effect of crop price and climate variables upon non-irrigated U.S. corn and soybean yields and acreage using a large county-level panel dataset from the 1977 to 2007 period. Their results indicated that climate change was attributable to declines in corn production ranging from 7 to 41 percent and declines in soybean production ranging from 8 to 45 percent when controlling for price effects. They also found that omitting price variables resulted in an overestimation of the climate change impact upon corn yields by up to 9 percent and upon soybean yields by up to 15 percent.

Burke and Emerick (2016) examined the impact of recent large variations in temperature and precipitation upon long-run adaption strategies to climate change by U.S. producers. Their results suggested that farmers were no more able to mitigate the negative impacts of climate change in the long-run as compared to in the short-run. They provide evidence that the lack of adaption was not driven by a lack of awareness regarding climate change rather they attribute this observation to the fact that farmers either lacked adequate long-run adaption options or they found those options available to be too expensive to implement. Using climate projections from

18 different climate change models, the authors projected that annual U.S. corn productivity would decline by roughly 15 percent by the year 2050, which is on par with the losses experienced during the severe drought of 2012.

Haile et al. (2017) analyzed determinants of global crop production for corn, wheat, rice, and soybeans over the period of 1961 to 2013 using seasonal production, price change, and price volatility data at the country level along with climate data from 32 global circulation models. Their results indicated that price and weather extremes not only have adverse impacts upon global food production but they also positively contribute to year-to-year fluctuations in global food availability. Using the climate predictions from the 32 GCM's, they project that climate change will reduce global food production by up to 9 percent in the 2030s and by 23% in the 2050s with a large heterogeneity across countries and crops. Their results also indicate that improvements in technology and agronomic practices have the capacity to offset some of the negative consequences of climate change impacts upon variability in food availability along with production.

Fei, McCarl and Thayer (2017) examined the effects of historical patterns in precipitation, temperature, and atmospheric gases along with the frequency of extreme weather events upon acreage adaption of cereal grains with a focus upon the Pacific Northwest region. Their results indicate that under climate change, in general, wheat production shifts northward in the Southern Great Plains, westward in the Northern Great Plains, and eastward in Oregon and Washington which are all cooler climates. They also found that overall wheat production would decline from 6 million acres under the no climate change scenario to between 5.4 and 5.7 million acres under the four climate change scenarios that were examined. Production declines in Oregon and Washington were partially offset by increases in Idaho when comparing the no change to the climate change scenarios. They also found that winter wheat would supplant spring wheat varieties along the northern border as an adaption to the warmer climate.

Li, Miao and Khanna (2019) examined the expansion of ethanol production in the United States and its impact upon land-use when controlling for the effect of changes in relative crop prices. Their empirical analysis utilized county level acreage data for the 2003 to 2014 crop years. They estimated a reduced-form, two-equation econometric model where the dependent variables were county level corn acreage and aggregate crop acreage. The explanatory variables included state-level corn prices, an aggregated crop price index, a fertilizer price index, population density, and average precipitation. Ethanol capacity and prices were modeled as endogenous to the system using a panel data instrumental variable estimator with county fixed effects. Ethanol capacity was modeled using an interaction term between railroad density associated with the county and the volume of ethanol mandated under the Renewable Fuel Standard (RFS). Prices were modeled using lagged crop stocks, and natural gas prices as a proxy for fertilizer costs. They found that an increase in ethanol capacity has led to a modest 3% increase in corn acreage and less than a 1% increase in total crop acreage between 2008 and 2012. Prices were found to have effects twice as large as ethanol capacity but this effect was essentially reversed by the sharp downturn in prices after 2012.

DATA AND METHODOLOGY

To calculate the relative over- or under-performance of annual corn and soybean production, a *production performance index* (PPI) was calculated as a proxy for the state and national production level relative to recent history. The formula for the PPI is as follows:

$$PPI_t = P_t - O(P_{t-1}, \cdots, P_{t-5}), \tag{1}$$

where P_t is the production level in time period t, and $O(\cdot)$ is the Olympic average function (drop minimum and maximum values and average the remaining three values). Therefore, the PPI measures the degree by which the current year's production either exceeded (over-performed) or fell short of (under-performed) the normal production level from the preceding five years. The Olympic average was used to minimize the impact of any extremely good or bad years when setting the past benchmark of what could be considered a "normal" production level for a particular region (state or national) at the time of comparison.

The PPI is illustrated for the U.S. and Iowa in Figure 2. As would be expected, there is a high degree of correlation (87.5 percent) between the two measures since Iowa is a major source of U.S. corn production. However, the Iowa measure is less volatile than the U.S. measure due to its smaller area.

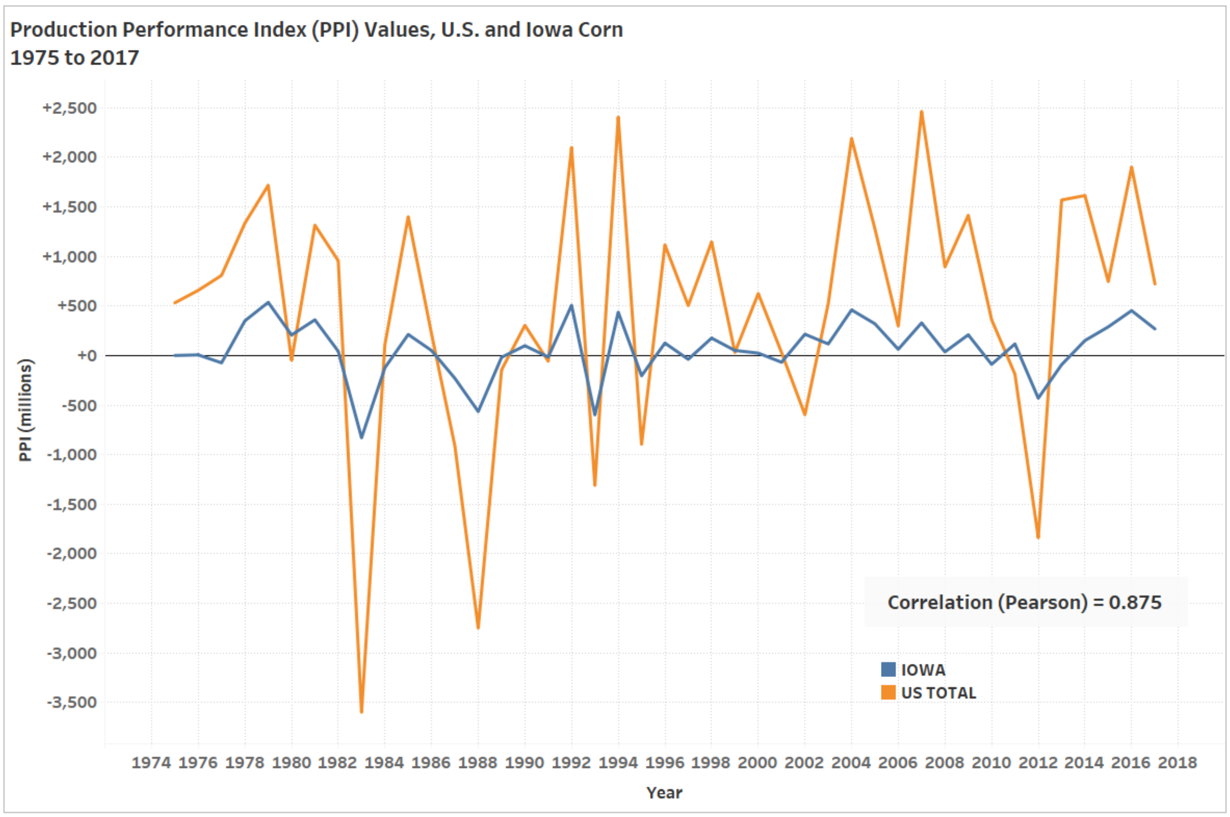

Figure 2. Production Performance Index (PPI) Values for U.S. and Iowa, 1975 to 2017

To analyze and rank the relative contribution of each state to the overall U.S. PPI index, a standard approach might be to estimate a linear regression model with the U.S. PPI as the dependent variable and the individual states' PPI's as the independent variables. The standardized coefficient values from the linear regression could be used to rank the individual states and test for significance. However, there are a couple of issues that arise with the application of linear regression to this particular dataset: (1) with the dataset split between 1975 to 1995 and 1996 to 2017, there are nearly as many independent variables as there are observations (referred to as the *sparsity* problem), and (2) the independent variables have a high level of correlation (*multicollinearity*) which can overstate the variance of the individual estimators and result in inaccurate individual coefficient values and t-statistics (Kennedy 1998).

Traditional econometric approaches to sparse data and multicollinearity typically involve one of four approaches (Kuhn and Johnson 2013): (1) *acknowledge and ignore* the problem, (2) *leave out or combine* the most problematic variables, (3) utilize a *penalized regression method* such as *ridge* or *lasso* regression, or (4) utilize a regression method based upon principal components such as *principal components regression* (PCR) or *partial least squares* (PLS). Another, more recent development in this area has been the introduction (Magidson 2010) of *correlated component regression* (CCR) which has some similarities but also some major differences with PLS.

Using a penalized or principal component based method typically involves the trade-off of increasing bias in exchange for a reduced variance in order to reduce the overall mean squared error (MSE) of the model prediction. In the case of penalized regression models, this is accomplished by adding a penalty function to the least squares optimization procedure that will tend to shrink the coefficient estimates of some variables towards zero in order to reduce the penalty cost. While generally effective at variable reduction when compared to stepwise regression and other variable reduction techniques, penalized models have the disadvantage of ignoring the presence of *suppressor variables* in the dataset. These unobserved, latent variables are not directly measured in the independent variable set but may have an important bearing upon the predictive ability of the model by enhancing the prediction capability of the visible, observed independent variables.

In the case of principal components regression (PCR), the unobserved components are derived from a subset of the eigenvalues of the variance-covariance or correlation matrix of the independent variable set. The number of eigenvalues utilized is generally chosen by setting a percentage threshold on the amount of total variable explained by the eigenvalues or by visual examination of the eigenvalues using a *scree plot*. The eigenvectors (principal components) from the chosen set of eigenvalues have the advantage of being orthogonal (have zero correlation). However, the principal components are optimized using information only from the set of independent variables and do not consider information contained in the dependent variable. Therefore, one may come up with cases where regressing the dependent variable upon the principal components results in significant information from the underlying independent variable set being disregarded through the elimination of lower ranked eigenvalues or through the removal of principal components that contain useful information but have statistically insignificant regression coefficients.

The use of *supervised PCR* (Bair et al. 2006) provides some improvement since each independent variable is assured of having its predictive significance evaluated (provided the matrix of independent variables is fully identified and non-sparse so that the OLS estimation procedure does not break down). However, supervised PCR still has the disadvantage of overlooking components that may act as suppressor variables.

Partial least squares (PLS) overcomes both of the aforementioned issues with PCR since the dependent variable is included with the independent variables in deriving the orthogonal components such that the optimization of each vector is heavily weighted in the direction of the dependent variable through correlation scores within an iterative weighting process. Additionally, the PLS regression procedure takes into account the impact of unseen latent suppressor variables. However, for explanatory modeling, PLS has a significant drawback in that the procedure will naturally be biased in the direction of predictors with the highest variance. This requires that the independent predictor variables be preprocessed via normalization and can result in over- and under-statement of the standardized regression coefficients. Also, due to the preprocessing, it is more difficult to quantify the relationship of each independent variable to the identified latent suppressor variables which makes their identification more difficult.

The correlated component regression (CCR) model is based upon two tuning parameters: the number of components to be derived (k) and the number of independent variables to retain in the model (p). Tuning of these parameters is done using a cross-validation procedure such as *m-fold validation*. A cross-validation metric such as R^2, mean squared error (MSE), or area under the receiver operating curve (AUROC) is chosen for optimization under the cross-validation procedure. The CCR procedure is extremely flexible and versions have been developed for modeling ordinary least squares (CCR-Linear), logistic (CCR-Logistic), linear discriminant (CCR-LDA), survival (CCR-Cox), and latent variable (CCR-Latent) models. For CCR-Logistic and CCR-LDA, the dependent variable is limited to two values (binary) so it cannot be applied to multinomial models.

The CCR-Linear algorithm begins with all P of the independent variables and estimates the following P single-variable regression equations:

$$Y_i = \delta_g^{(1)} + \lambda_g^{(1)} \cdot X_{g,i}, \tag{2}$$

where Y_i is observation i of the dependent variable with $i = 1,\ldots,N$; $X_{g,i}$ is observation i of independent variable $g = 1,\ldots,P$; $\delta_g^{(1)}$ and $\lambda_g^{(1)}$ are the regression intercept and slope parameters for independent variable g.

The first correlated component variable, CC_1, is then constructed as the linear, weighted average of each predictor using the single equation slope coefficients as the weights:

$$CC_{1,i} = \frac{1}{P} \cdot \sum_{g=1}^{P} \hat{\lambda}_g^{(1)} \cdot X_{g,i}, \tag{3}$$

From (3), the 1-component model is then estimated as:

$$Y_i = \alpha^{(1)} + \beta_1^{(1)} \cdot CC_{1,i}, \tag{4}$$

with the relevant cross-validation (CV) metric (CV-R^2 or CV-MSE) stored from this regression model for later determination of the optimal number of retained components in the model. The CC_1 component is called the *direct effects* component since it measures the direct impact of each independent variable upon the dependent variable without any latent suppressor effects.

The second component variable (CC_2) is constructed by first estimating the following regression equation for each of the independent variables:

$$Y_i = \delta_g^{(2)} + \gamma_{1,g}^{(2)} \cdot CC_{1,i} + \lambda_g^{(2)} \cdot X_{g,i}, \tag{5}$$

then, using the results from (5) to derive the correlated component:

$$CC_{2,i} = \frac{1}{P} \cdot \sum_{g=1}^{P} \hat{\lambda}_g^{(2)} \cdot X_{g,i}, \tag{6}$$

The second component CV metric is then derived from the following regression equation:

$$Y_i = \alpha^{(2)} + \beta_1^{(2)} \cdot CC_{1,i} + \beta_2^{(2)} \cdot CC_{2,i}, \tag{7}$$

Note that CC_2 and subsequent derived component variables represent *latent suppressor effect* variables in the correlated component model.

The progression of component derivations can continue up to when the number of components equals the number of independent variables if the model is fully identified and non-sparse, or the maximum number specified by the user. In the case where the number of components equals the number of independent variables, the CCR model will be equivalent to the OLS regression model. The initial regression model would be:

$$Y_i = \delta_g^{(P)} + \gamma_{1,g}^{(P)} \cdot CC_{1,i} + \cdots + \gamma_{P-1,g}^{(P)} \cdot CC_{P-1,i} + \lambda_g^{(P)} \cdot X_{g,i}, \tag{8}$$

With the final component equal to:

$$CC_{P,i} = \frac{1}{P} \cdot \sum_{g=1}^{P} \hat{\lambda}_g^{(p)} \cdot X_{g,i}, \tag{9}$$

and

$$Y_i = \alpha^{(P)} + \beta_1^{(P)} \cdot CC_{1,i} + \cdots + \beta_P^{(P)} \cdot CC_{P,i}, \tag{10}$$

to determine the final cross-validation metric.

Once the optimal number of components (K^*) is determined using the optimal value of the cross-validation (CV) metric, the number of independent variables can either be specified by the user (up to P) or can be optimally determined using a *step-down procedure*. The first step in the step-down procedure is to estimate the model with all of the predictors and calculate the cross-validation metric. In the next step, the independent variable with the smallest absolute value of its standardized coefficient is removed and the model is re-estimated with the reduced set of independent variables. This is repeated until there is only one independent variable left in the dataset. The optimal set of independent variables (P^*) is the set that produces the optimal value of the cross-validation metric.

Initial results presented by Magidson (2010) indicated that CCR can perform as well or better than PCA-based and penalty function approaches when forecasting out-of-sample values from sparse and multicollinear datasets. In another paper (Magidson and Wassmann 2010), the potential value of latent suppressor variables in the detection of prostate cancer was demonstrated using an application of CCR-LDA to patient data.

To date, most published applications of CCR have been in the medical and sociological fields (Alkerwi et al. 2015). Given its recent development, applications of CCR in the economics literature have been relatively sparse with the exception of a paper presented at a math and engineering conference (Trivedi and Birau 2013) that examined correlation between several international stock indices and a recent logistics paper (Garver and Williams 2018).

The dataset is comprised of the state-level corn and soybean production from 1970 to 2017 for the 18 major corn producing (Colorado, Illinois, Indiana, Iowa, Kansas, Kentucky, Michigan, Minnesota, Nebraska, North Carolina, North Dakota, Ohio, Pennsylvania, South Dakota, Tennessee, Texas, and Wisconsin) and soybean producing (Arkansas, Illinois, Indiana, Iowa, Kansas, Kentucky, Louisiana, Michigan, Minnesota, Mississippi, Missouri, Nebraska, North Carolina, North Dakota, Ohio, South Dakota, Tennessee, and Wisconsin) states as determined by the USDA. The dataset also contains the national-level production for both crops and an “Other States” aggregate is derived as a residual when subtracting the sum of the 18-state production from the national aggregate. The data come from the USDA-NASS Quick Stats online database.[4] The 18-major states typically[5] comprise between 92.3 to 96.3 percent of the national corn and between 89.5 and 99.1 percent of the national soybean production over the time period examined.

To characterize and interpret the estimated correlated component (CC) variables, a Spearman’s rank-order correlation was calculated between each component and a set of corn production and weather data. The corn production data included yield versus trend (using trendline fit from 1948 to 2017), the difference between final planted acreage and the March 1st estimate from the USDA’s Prospective Plantings report (representing a proxy for prevented plant acres), and an excess acreage abandonment proxy measured by the difference between the percent of planted acres not harvested and the Olympic average of the same measure over the previous five years.

The weather variables come from data in the U.S. National Weather Service’s Oceanic Niño Index (ONI)[6], version 5, and a set of annual weather measurements for Des Moines, Iowa covering daily temperatures (average, maximum, minimum, extreme maximum, and extreme minimum), days with minimum temperatures below a threshold (zero, 32 degrees Fahrenheit), days with maximum temperatures above a threshold (70 and 90 degrees Fahrenheit), heating degree days, cooling degree days, precipitation (total and extreme maximum daily in inches), days with precipitation above 0.1 inches and 1.0 inches, snowfall (total and extreme daily maximum), number of days with snow depth greater than one inch, and highest daily snow depth in inches. All of the Des Moines weather measurements come from data in the *NOAA Climate Data Online*[7] database.

---

[4] https://quickstats.nass.usda.gov

[5] Based upon 95% confidence interval.

[6] http://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php

[7] https://www.ncdc.noaa.gov/cdo-web/

RESULTS

In this paper, the Magidson CCR model was applied to the examination of the relative importance of each state in predicting the PPI on a national scale for both corn and soybeans. This was done over two separate time periods (1) the 1975 through 1995 crop years and (2) the 1996 through 2017 crop years. The state-level rankings and impact are derived from the standardized coefficients of the CCR regression procedure. For each crop and time period, a *Herfindahl-Hirschman Index* (HHI) was calculated by summing the squared percentage shares of the standardized coefficients (absolute value). This allowed for an examination of the degree and change in geographic concentration across the two time periods. A primary hypothesis examined was whether the introduction of new crop technologies combined with changes in climate and farm policy resulted in a greater geographic dispersion of the relative influence of each state upon the national aggregate. This would imply a decline in the HHI from the 1975-1995 to the 1996-2017 periods for each crop.

The results presented in this paper were derived using the COREXpress™ software package from Statistical Innovations (www.statisticalinnovations.com). The maximum number of correlated components (p) was set to eight for each estimation and the step-down procedure of variable selection was not utilized (all of the states were retained in the estimated model). Each model was estimated using a 4-fold cross-validation option using R^2 as the CV metric in the CCR-Linear procedure.

U.S. Corn: Initial Time Period (1975-1995)

The cross-validation procedure for corn in the pre-GMO time frame (1975 to 1995) resulted in a maximum CV-R^2 of 0.987 with 4 components retained. The estimated correlated component coefficient values ($\hat{\beta}_i$) are shown in the second column of Table 1 along with the coefficient standard errors, t-statistics, and levels of significance. All four components are highly significant with p-values less than 0.01. The standardized coefficient values and the percentage shares are also shown in the last two columns of the table. The direct effects component (CC_1) has almost a 2/3 (65.1%) share of the total standardized coefficient value. The next two indirect effects components (CC_2 and CC_3) have an almost equal share (13.7% and 14.3%) while the final component (CC_4) has the lowest share at 6.9 percent. Note that the percent shares of the standardized totals can be used as a proxy for the percent of variability explained by the component.

Table 1. Correlated Component Regression and Standardized Coefficients, Corn, 1975 to 1995 Period

Correlated Component	Value	Standard Error	T-Statistic	Pr > |t|	Signif*	Standardized Value	Share of Standardized Total (%)
CC_1	0.0754	0.0027	28.327	< 0.0001	***	0.8150	65.1%
CC_2	0.2861	0.0201	14.251	< 0.0001	***	0.1715	13.7%
CC_3	0.4415	0.0451	9.780	< 0.0001	***	0.1786	14.3%
CC_4	0.5861	0.1527	3.839	0.0014	***	0.0865	6.9%

*Three-stars = signficantly different from zero at 99% level, two-stars = 95%, one-star = 90%.

The vector product of each states' (X_g for g = 1,…, 19) component loading ($\hat{\lambda}_g^{(i)}$; i = 1,…, 4) with its individual PPI value in each year to derive the correlated component values by year (t =

1975,…,1995) and these values can be plotted to provide additional information about the characterization of each component. Figure 1 shows the values of the four correlated components for corn by year during the 1975-1995 period.

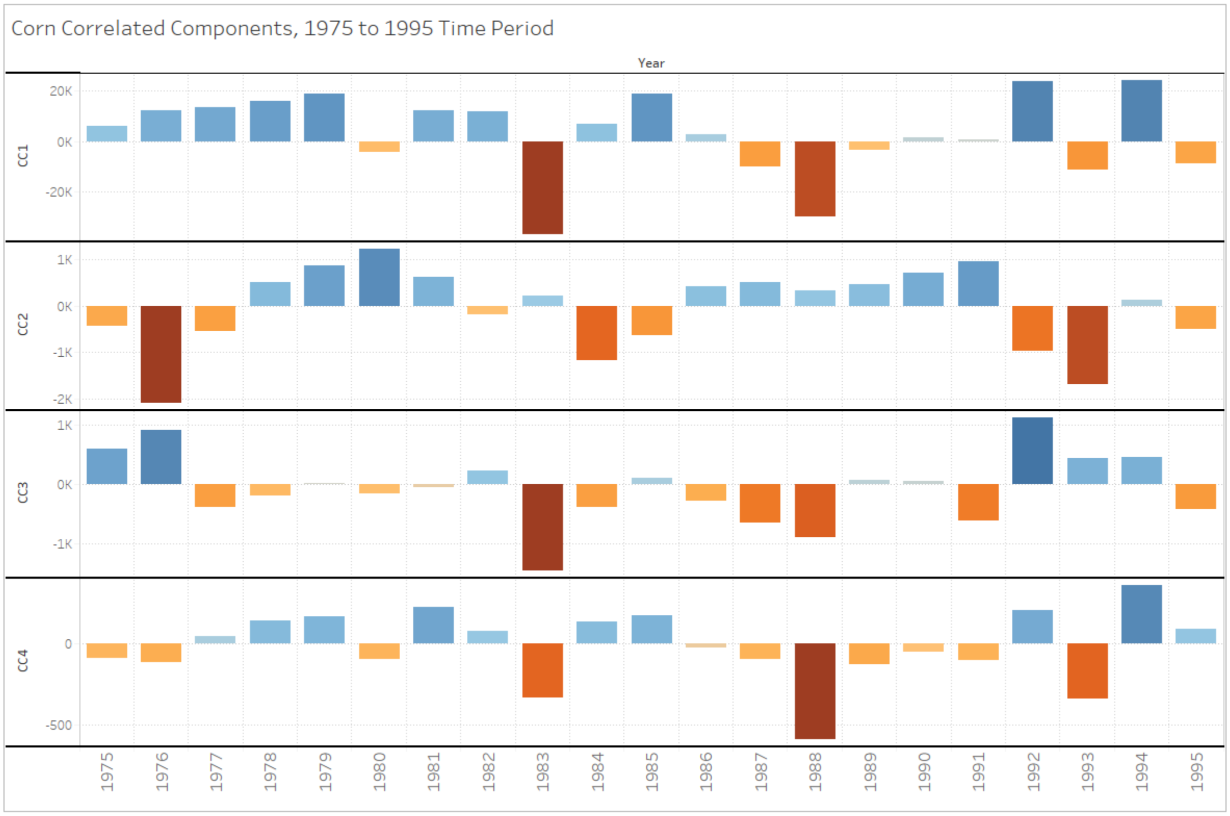

Figure 3. Correlated Component Values for Corn, 1975 to 1995 Period

As expected, the direct effects component (CC_1) is highly correlated with the corn production metrics (yield versus trend, excess abandonment, and planted acreage differential) with signs in the expected directions for all three (positive for yield versus trend and planted acreage differential, and negative for excess abandonment). The first indirect effects component (CC_2) is positively correlated with cooling degree days, average daily temperature, and average minimum daily temperature – reflecting a suppressor effect tied to cooler than normal years such as 1976 and 1993. The second indirect effects component (CC_3) is negatively correlated with cooling degree days, the extreme maximum temperature, days with maximum over 90 degrees, average minimum temperature, average temperature, and yield versus trend – reflecting a suppressor effect tied to very hot years (such as 1983 and 1988). The final indirect effects component (CC_4) is negatively correlated with excess abandonment and cooling degree days while positively correlated with yield versus trend and extreme maximum snowfall – likely reflecting a suppression effect tied to late season conditions and abandonment.

Table 2. Individual States' Coefficient Values and Shares, Corn, 1975 to 1995 Period

Rank*	State	Coefficient	Standard Error	T-Statistic	Pr > |t|	Signif***	Standardized Coefficient	Share(%)*
1	IOWA	1.167	0.061	19.225	< 0.0001	***	0.282	23.1%
2	ILLINOIS	0.993	0.053	18.659	< 0.0001	***	0.201	16.5%
3	MINNESOTA	0.810	0.074	10.990	< 0.0001	***	0.092	7.6%
4	INDIANA	0.899	0.071	12.702	< 0.0001	***	0.088	7.2%
5	SOUTH DAKOTA	1.577	0.122	12.896	< 0.0001	***	0.068	5.6%
6	MICHIGAN	2.091	0.126	16.641	< 0.0001	***	0.068	5.6%
7	KANSAS	2.373	0.092	25.890	< 0.0001	***	0.062	5.1%
8	WISCONSIN	1.050	0.074	14.236	< 0.0001	***	0.060	4.9%
9	OTHER STATES	0.952	0.052	18.193	< 0.0001	***	0.058	4.8%
10	MISSOURI	0.995	0.053	18.823	< 0.0001	***	0.044	3.6%
11	NORTH CAROLINA	2.197	0.115	19.028	< 0.0001	***	0.039	3.2%
12	OHIO	0.604	0.072	8.356	< 0.0001	***	0.038	3.1%
13	KENTUCKY	1.341	0.227	5.911	< 0.0001	***	0.031	2.5%
14	NORTH DAKOTA	3.313	0.727	4.556	0.0003	***	0.025	2.0%
15	PENNSYLVANIA	1.243	0.355	3.496	0.0030	***	0.024	2.0%
16	NEBRASKA	0.126	0.082	1.532	0.1450		0.012	1.0%
17	COLORADO	0.835	0.367	2.277	0.0369	**	0.010	0.9%
18	TENNESSEE	0.866	0.875	0.990	0.3369		0.010	0.8%
19	TEXAS	-0.184	0.107	-1.714	0.1059		-0.005	0.4%
	[Constant]	23.144	19.423	1.192	0.2508			
Herfindahl-Hirschman Index**								**1,103**

*Rank and share based upon absolute value of standardized coefficient.

**Equals sum of squared percentage shares (range from 0 to 10,000).

***Three-stars = signficantly different from zero at 99% level, two-stars = 95%, one-star = 90%.

The individual states' non-standardized and standardized coefficient values along with the percent standardized share (absolute value) are shown in Table 2. The states are ranked in order by their share of the absolute standardized coefficient values. Note that each state's (g) non-standardized coefficient value ($\hat{\phi}_g$) is calculated as:

$$\hat{\phi}_g = \sum_{i=1}^{4} \hat{\lambda}_g^{(i)} \cdot \hat{\beta}_i, \tag{11}$$

One of the weaknesses of the COREXpress software package is that it does not provide coefficient standard errors and t-statistics for the regression coefficients; however, these can be derived from the regression standard errors of the component coefficients as follows:

$$se(\hat{\phi}_g) = \sqrt{\sum_{i=1}^{4} \left(\hat{\lambda}_g^{(i)}\right)^2 \cdot \left(se(\hat{\beta}_i)\right)^2}, \tag{12}$$

where $se(\cdot)$ is the standard error for the regression coefficient. The constant coefficient represents the intercept ($\hat{\alpha}$) from the optimal CV-R^2 regression.

The results indicate that during the 1975 to 1995 period, Iowa and Illinois had a dominant share (39.6%) of the influence upon the overall national corn production outcome. The next tier of states (Minnesota and Indiana) had an approximate 15% share. Note all of the coefficients are statistically significant at the 95% confidence level with the exception of Nebraska, Tennessee, and Texas. The Herfindahl-Hirschman Index of 1,103 is below the 1,500 threshold that the U.S.

Department of Justice would consider a moderate level of concentration and well below the 2,500 threshold that would be considered highly concentrated.

Table 3 shows each states' standardized loading upon each of the four correlated components. The cells are shaded based upon the coefficient ranking for each component with light yellow to green representing positive loadings and dark yellow to red representing negative loadings. Iowa and Illinois both have positive loadings on all four components. Minnesota has a negative loading on the high temperature component (CC_3) while Indiana has a negative loading upon the cool temperature component (CC_2). Nebraska, with its heavier influence upon irrigated production, is less susceptible to the high temperature (CC_3) and the abandonment (CC_4) components, having a strong negative loading on both. The same holds true for Colorado to a lesser extent.

Table 3. Individual States' Standardized Loadings on Correlated Components, Corn, 1975 to 1995 Period

State	Correlated Component			
	CC1	CC2	CC3	CC4
IOWA	0.087	0.464	0.547	0.390
ILLINOIS	0.086	0.131	0.505	0.208
MINNESOTA	0.076	0.185	-0.162	0.319
INDIANA	0.077	-0.122	0.316	-0.122
SOUTH DAKOTA	0.073	0.145	-0.165	0.150
MICHIGAN	0.074	0.089	0.028	-0.146
KANSAS	0.050	0.008	0.091	0.041
WISCONSIN	0.070	0.115	-0.143	0.100
OTHER STATES	0.070	-0.104	0.116	-0.017
MISSOURI	0.068	0.007	-0.054	-0.036
NORTH CAROLINA	0.057	-0.049	0.026	-0.043
OHIO	0.079	-0.160	0.074	-0.141
KENTUCKY	0.079	-0.332	0.135	-0.011
NORTH DAKOTA	0.052	0.011	-0.193	0.173
PENNSYLVANIA	0.076	-0.277	-0.066	0.249
NEBRASKA	0.079	0.078	-0.235	-0.272
COLORADO	0.031	0.064	-0.054	-0.190
TENNESSEE	0.067	-0.290	-0.154	0.370
TEXAS	0.032	-0.156	0.020	-0.091

U.S. Corn: Latter Time Period (1996-2017)

The four-fold cross-validation procedure resulted in the optimal selection of five components in the post-GMO period (1996 to 2017). The CV-R^2 value was 0.9733. The component regression coefficients (non-standardized and standardized) are shown in Table 4. All five components are significant at the 99% confidence level. The direct effects (CC_1) coefficient has just under a 2/3 share (60.7%) of the standardized value with the remaining 39.3% in the indirect effects components (CC_2, CC_3, CC_4, and CC_5).

Table 4. Correlated Component Regression and Standardized Coefficients, Corn, 1996 to 2017 Period

Correlated Component	Value	Standard Error	T-Statistic	Pr > |t|	Signif*	Standardized Value	Share of Standardized Total (%)
CC_1	0.1580	0.0033	47.6970	< 0.0001	***	0.985	60.7%
CC_2	0.3938	0.0227	17.3549	< 0.0001	***	0.270	16.6%
CC_3	0.5297	0.0470	11.2639	< 0.0001	***	0.197	12.2%
CC_4	0.7200	0.1381	5.2132	< 0.0001	***	0.106	6.5%
CC_5	0.7062	0.2335	3.0246	0.0081	***	0.064	4.0%

*Three-stars = signficantly different from zero at 99% level, two-stars = 95%, one-star = 90%.

Figure 4 shows the plot of the component values by year. An examination of the Spearman rank-order correlation coefficients with the production and weather metrics indicates that the direct effects coefficient (CC_1) has significant (p-value of 0.1 or less) negative correlation with the levels of excess abandonment, extreme maximum temperature, number of days with maximum temperatures over 90 degrees. CC_1 has significant positive correlations with yield versus trend, days with minimum temperature below zero, days with maximum temperature less than 32 degrees, total precipitation, days with precipitation greater than 1 inch, and days with precipitation greater than 0.1 inch. The plot shows that CC_1 has mostly positive to neutral values for most years with the exception of 2002 and the extreme drought year of 2012.

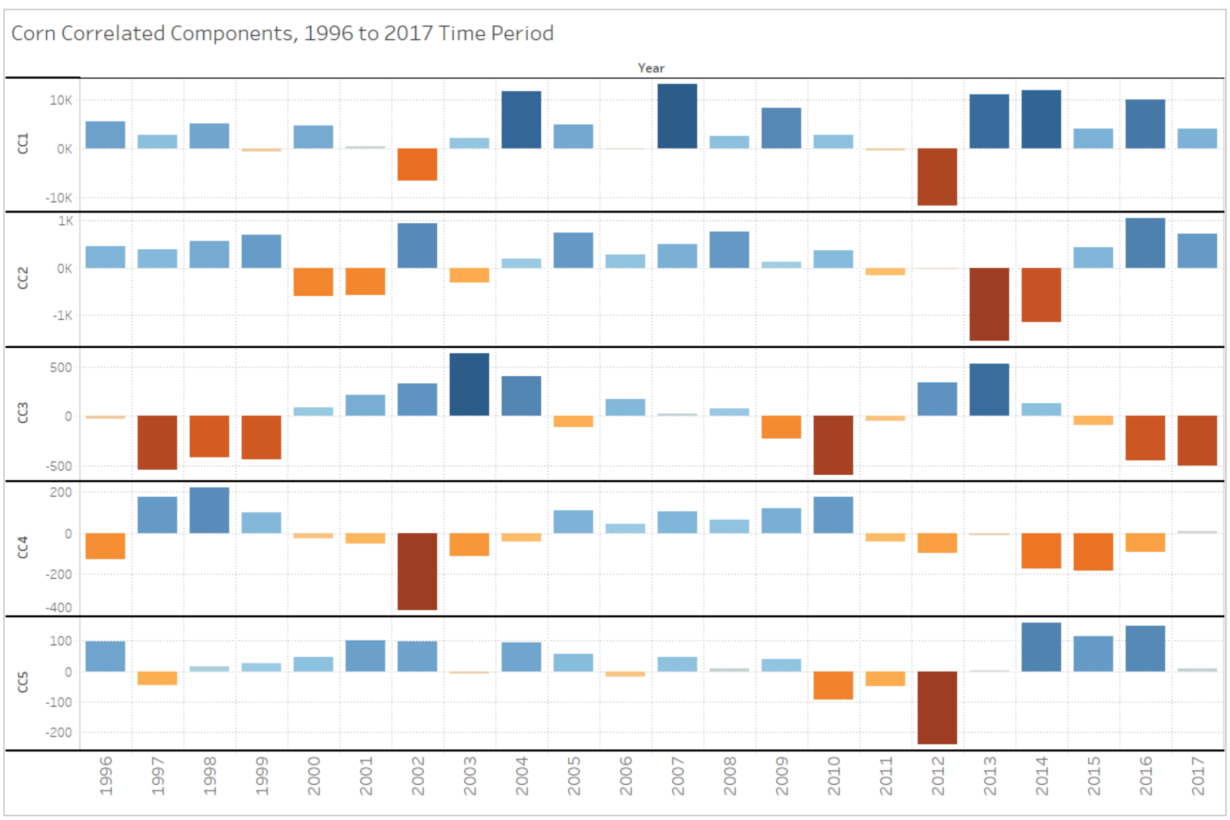

Figure 4. Correlated Component Values for Corn, 1996 to 2017 Period

The first indirect effects component (CC_2) has a significant negative correlation on the change in the annual average ONI index and a positive correlation with the average temperature level. Strong positive increases in the annual average ONI are associated with strong El Niño years (1997 and 2015) indicating that this component reflects a suppressor effect tied to the El Niño cycle. The second indirect effects component (CC_3) has significant positive correlations with the extreme minimum temperature, the 2nd half (last six months of calendar year) ONI index value, and days with a minimum temperature less than 32 degrees. This coefficient also is negatively correlated with the extreme maximum temperature, the average daily minimum temperature, the total level of precipitation, and days with precipitation greater than 1.0 inch. This would indicate a suppression effect tied to hot, humid conditions – particularly in the latter half of the growing season. The third indirect effects component (CC_4) has a strong positive correlation with the 2nd half average ONI and days with precipitation greater than 1.0 inch. It also has positive correlation with total snowfall and days with highest daily snow depth greater than or equal to 1 inch. This component is likely a suppressor effect tied to the overall level of soil moisture and precipitation. The fourth and final indirect effects coefficient (CC_5) has a significant negative correlation with the extreme maximum temperature and a positive correlation on yield versus trend. It also has a negative correlation on cooling degree days, days with maximum temperature above 90 degrees, and excess abandonment. This component obviously represents a suppressor effect related to extremely hot conditions – particularly later in the growing season.

Table 5. Individual States' Coefficient Values and Shares, Corn, 1996 to 2017 Period

Rank*	State	Coefficient	Standard Error	T-Statistic	Pr > |t|	Signif***	Standardized Coefficient	Share(%)*
1	ILLINOIS	0.874	0.022	39.089	< 0.0001	***	0.228	13.5%
2	NEBRASKA	1.540	0.048	32.145	< 0.0001	***	0.215	12.8%
3	IOWA	0.927	0.043	21.724	< 0.0001	***	0.194	11.5%
4	INDIANA	1.526	0.107	14.267	< 0.0001	***	0.181	10.7%
5	KANSAS	1.393	0.180	7.738	< 0.0001	***	0.133	7.9%
6	SOUTH DAKOTA	1.066	0.031	34.041	< 0.0001	***	0.097	5.8%
7	MISSOURI	0.877	0.112	7.837	< 0.0001	***	0.079	4.7%
8	OTHER STATES	0.668	0.040	16.525	< 0.0001	***	0.076	4.5%
9	PENNSYLVANIA	2.694	0.181	14.853	< 0.0001	***	0.070	4.2%
10	TEXAS	1.202	0.121	9.944	< 0.0001	***	0.061	3.6%
11	NORTH CAROLINA	3.085	0.154	19.994	< 0.0001	***	0.059	3.5%
12	COLORADO	-2.176	0.522	-4.167	0.0007	***	-0.053	3.1%
13	MINNESOTA	0.523	0.027	19.321	< 0.0001	***	0.051	3.1%
14	TENNESSEE	-2.887	0.587	-4.920	0.0002	***	-0.049	2.9%
15	NORTH DAKOTA	0.706	0.033	21.116	< 0.0001	***	0.043	2.5%
16	WISCONSIN	0.784	0.099	7.912	< 0.0001	***	0.037	2.2%
17	MICHIGAN	1.123	0.119	9.461	< 0.0001	***	0.029	1.7%
18	KENTUCKY	0.488	0.231	2.111	0.0509	*	0.015	0.9%
19	OHIO	0.178	0.149	1.193	0.2503		0.014	0.8%
	[Constant]	84.279	15.860	5.314	0.0001	***		
Herfindahl-Hirschman Index**								**818**

*Rank and share based upon absolute value of standardized coefficient.

**Equals sum of squared percentage shares (range from 0 to 10,000).

***Three-stars = signficantly different from zero at 99% level, two-stars = 95%, one-star = 90%.

Table 5 shows the individual states' coefficients (non-standardized and standardized) ranked by their absolute percent share of the total standardized values. Almost half (48.5%) of the total value is in the top four states (Illinois, Nebraska, Iowa, and Indiana). All of the coefficients are statistically significant at the 90% level or higher except for Ohio which is also the lowest ranked among the states. The Herfindahl-Hirschman Index is well below the DOJ guidelines for a concentrated market.

Table 6. Individual States' Standardized Loadings on Correlated Components, Corn, 1996 to 2017 Period

	Correlated Component				
State	CC1	CC2	CC3	CC4	CC5
ILLINOIS	0.133	0.180	0.242	-0.013	0.038
NEBRASKA	0.129	0.208	0.055	0.242	-0.069
IOWA	0.113	0.275	0.154	-0.304	0.152
INDIANA	0.122	-0.128	0.129	0.407	0.412
KANSAS	0.117	0.152	-0.515	0.418	0.535
SOUTH DAKOTA	0.116	-0.082	0.012	0.039	-0.020
MISSOURI	0.127	-0.083	-0.077	-0.296	0.350
OTHER STATES	0.061	-0.021	0.143	0.040	-0.168
PENNSYLVANIA	0.073	-0.093	0.147	0.037	-0.154
TEXAS	0.103	0.053	-0.120	-0.179	-0.185
NORTH CAROLINA	0.058	-0.054	0.126	-0.056	-0.033
COLORADO	0.037	-0.012	-0.295	0.066	-0.539
MINNESOTA	0.012	0.168	-0.028	-0.001	-0.003
TENNESSEE	0.058	-0.232	-0.028	-0.432	0.125
NORTH DAKOTA	0.015	0.056	0.077	0.015	-0.054
WISCONSIN	0.057	0.061	-0.229	0.043	0.078
MICHIGAN	0.016	-0.029	0.078	0.101	-0.081
KENTUCKY	0.106	-0.287	0.024	-0.227	0.116
OHIO	0.091	-0.269	-0.134	0.417	-0.315

Each states' standardized loadings on each component is shown in Table 6. Note that, unlike the pre-GMO period, every state has at least one negative loading upon at least one of the correlated components.

U.S. Corn: Summary of Changes in State Rankings and Concentration

A summary of the change in states' rankings and standardized shares along with the change in the HHI is contained in Table 7. In terms of rankings, the states with the largest increases are Nebraska (from 16th to 2nd), Texas (19th to 10th), Pennsylvania (15th to 9th), Colorado (17th to 12th), and Tennessee (18th to 14th). Note that three of these states (Nebraska, Texas, and Colorado) have significant acreage under irrigation. States with the largest declines are Michigan (6th to 17th), Minnesota (3rd to 13th), Wisconsin (8th to 16th), Ohio (12th to 19th), and Kentucky (13th to 18th). Three of these states (Michigan, Minnesota, and Wisconsin) are among the northernmost tier of the Corn Belt. In terms of percentage share, Nebraska (+11.8%) had the largest gain while Iowa (-11.6%) had the largest decline.

Table 7. Change in State-Level Rankings and Shares, Corn

State	Rank			Share		
	1975 to 1995	1996 to 2017	Change	1975 to 1995	1996 to 2017	Change
IOWA	1	3	-2	23.1%	11.5%	-11.6%
ILLINOIS	2	1	+1	16.5%	13.5%	-3.0%
MINNESOTA	3	13	-10	7.6%	3.1%	-4.5%
INDIANA	4	4	0	7.2%	10.7%	+3.5%
SOUTH DAKOTA	5	6	-1	5.6%	5.8%	+0.2%
MICHIGAN	6	17	-11	5.6%	1.7%	-3.9%
KANSAS	7	5	+2	5.1%	7.9%	+2.8%
WISCONSIN	8	16	-8	4.9%	2.2%	-2.8%
OTHER STATES	9	8	+1	4.8%	4.5%	-0.2%
MISSOURI	10	7	+3	3.6%	4.7%	+1.1%
NORTH CAROLINA	11	11	0	3.2%	3.5%	+0.3%
OHIO	12	19	-7	3.1%	0.8%	-2.3%
KENTUCKY	13	18	-5	2.5%	0.9%	-1.6%
NORTH DAKOTA	14	15	-1	2.0%	2.5%	+0.5%
PENNSYLVANIA	15	9	+6	2.0%	4.2%	+2.1%
NEBRASKA	16	2	+14	1.0%	12.8%	+11.8%
COLORADO	17	12	+5	0.9%	3.1%	+2.3%
TENNESSEE	18	14	+4	0.8%	2.9%	+2.1%
TEXAS	19	10	+9	0.4%	3.6%	+3.2%

Herfindahl-Hirschman Index	**1103**	**818**	**-285**

The Herfindahl-Hirschman Index (HHI) declined by 286 points between the two periods indicating a minor lowering (by HHI standards) of the share concentration among the states. Therefore, one can conclude that the recent technology, climate, and farm policy changes resulted in a moderate dispersion of the geographic influence of corn production upon the national aggregate when compared to the earlier 1975-1995 period. The change in rankings also indicated that irrigation (Nebraska, Colorado, and Texas) may be a primary factor in the shifts in geographic importance for national corn production. However, in general, there appears to be a pronounced shift from East to West in terms of geographic importance with major gains in states such as Nebraska and Texas with major losses in Iowa, Minnesota, and Michigan.

U.S. Soybeans: Initial Time Period (1975-1995)

The four-fold cross-validation procedure resulted in an optimal CV-R^2 of 98.32% with three correlated components for the 1975 to 1995. Table 8 shows the estimates for the non-standardized and standardized coefficients along with statistical significance indicators and standardized coefficient shares. All three coefficients are statistically significant at the 99% confidence level. The direct effects coefficient (CC_1) has almost ¾ (74.7%) of the total standardized value. The remaining shares are almost evenly distributed across the two indirect effects coefficients with CC_2 having a slightly higher share.

Table 8. Correlated Component Regression and Standardized Coefficients, Soybeans, 1975 to 1995 Period

Correlated Component	Value	Standard Error	T-Statistic	Pr > |t|	Signif*	Standardized Value	Share of Standardized Total (%)
CC_1	0.1163	0.0026	45.3784	< 0.0001	***	0.9596	74.7%
CC_2	0.3625	0.0474	7.6399	< 0.0001	***	0.1889	14.7%
CC_3	0.3192	0.0684	4.6653	0.0002	***	0.1363	10.6%

*Three-stars = significantly different from zero at 99% confidence level, two-stars = 95%, one-star = 90%.

A plot of the correlated components over time is shown in Figure 5. The direct effects coefficient (CC_1) has statistically significant (p-value less than 0.1) and positive Spearman rank-order correlation with yield versus trend, extreme maximum snowfall, number of days with snow depth greater than 1 inch, days with minimum temperature below zero, days with maximum temperature below 32 degrees, and heating degree days. The one significant negative coefficient was on the excess abandonment variable. This coefficient also had negative coefficients for the average annual temperature, days with temperature over 90 degrees, extreme maximum temperature, and average maximum temperature.

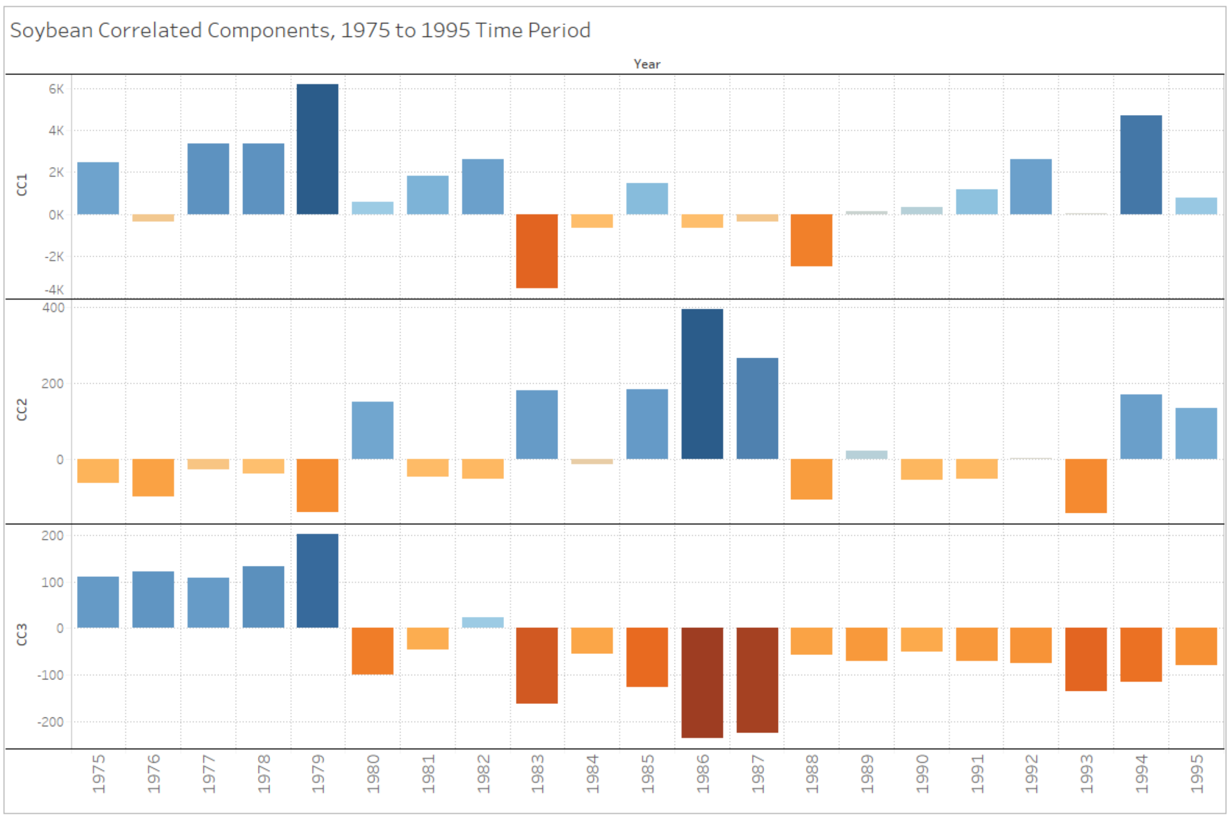

Figure 5. Correlated Component Values for Soybeans, 1975 to 1995 Period

The first indirect effects component (CC_2) did not have any statistically significant correlations with the production metric and weather variables. The highest correlation was a negative relationship with the days with minimum temperature below 32 degrees with had a p-value of 0.114. The next three ranked correlations (in level of significance) were all positive loadings on days with maximum temperature above 70 degrees, the extreme maximum temperature, and days with maximum temperature above 90 degrees. This indicates that the coefficient is positively influenced by hotter temperatures and negatively influenced by cooler temperatures.

The final indirect effects component (CC_3) had statistically significant positive relationship with days with minimum temperature below zero and negative correlation with 1st half average ONI and days with precipitation over 1.0 inch. The coefficient was also positively related to the planted acreage differential indicating a negative relationship with prevented plant acres. Overall, this indicates that the component reflects a sensitivity to early season conditions that impact the rate of planting for the crop.

Table 9 shows the state non-standardized and standardized coefficient values along with the statistical significance and standardized share indicators. All of the coefficients were statistically significant at the 99% level with the exception of Michigan which was significant at the 95% level.

Table 9. Individual States' Coefficient Values and Shares, Soybeans, 1975 to 1995 Period

Rank*	State	Coefficient	Standard Error	T-Statistic	Pr > |t|	Signif***	Standardized Coefficient	Share(%)*
1	MISSOURI	1.156	0.048	24.106	< 0.0001	***	0.140	9.8%
2	IOWA	0.818	0.037	21.934	< 0.0001	***	0.137	9.6%
3	ILLINOIS	0.743	0.024	31.443	< 0.0001	***	0.127	8.9%
4	MINNESOTA	1.169	0.055	21.324	< 0.0001	***	0.127	8.9%
5	OTHER STATES	0.705	0.067	10.458	< 0.0001	***	0.113	7.9%
6	OHIO	1.294	0.043	30.132	< 0.0001	***	0.106	7.4%
7	TENNESSEE	2.358	0.248	9.523	< 0.0001	***	0.102	7.1%
8	MISSISSIPPI	1.252	0.125	10.009	< 0.0001	***	0.083	5.8%
9	NEBRASKA	1.583	0.087	18.150	< 0.0001	***	0.081	5.6%
10	LOUISIANA	1.215	0.111	10.908	< 0.0001	***	0.069	4.9%
11	INDIANA	0.806	0.039	20.479	< 0.0001	***	0.066	4.6%
12	SOUTH DAKOTA	1.441	0.109	13.232	< 0.0001	***	0.051	3.6%
13	KANSAS	1.077	0.089	12.128	< 0.0001	***	0.048	3.3%
14	KENTUCKY	1.297	0.214	6.072	< 0.0001	***	0.046	3.2%
15	NORTH CAROLINA	1.511	0.146	10.375	< 0.0001	***	0.042	2.9%
16	ARKANSAS	0.414	0.064	6.509	< 0.0001	***	0.033	2.3%
17	WISCONSIN	1.652	0.141	11.707	< 0.0001	***	0.030	2.1%
18	NORTH DAKOTA	1.361	0.150	9.084	< 0.0001	***	0.018	1.3%
19	MICHIGAN	0.651	0.272	2.392	0.0286	**	0.012	0.8%
	[Constant]	-5.211	6.382	-0.816	0.4255			
Herfindahl-Hirschman Index**								**680**

*Rank and share based upon absolute value of standardized coefficient.

**Equals sum of squared percentage shares (range from 0 to 10,000).

***Three-stars = significantly different from zero at 99% confidence level, two-stars = 95%, one-star = 90%.

In terms of standardized share, there is not much distance between the top four states (Missouri, Iowa, Illinois, and Other States) ranging from 9.8 down to 8.9 percent. There is also not much distance separating the second tier of states (Ohio, Tennessee, and Mississippi) ranging from 7.9 to 7.1 percent. The Herfindahl-Hirschman Index value of 680 indicates a very low level of concentration among the shares and is much lower than the value for corn (1,103) in the same time period.

The individual states' standardized loadings upon the correlated components is contained in Table 10. Note the strong suppression effect for the Delta and southern Corn Belt states (Kentucky, Mississippi, Tennessee, and Arkansas) under CC2. The Plains states (Kansas, Nebraska, and South Dakota) have strong suppression effects under CC3 along with Michigan and Indiana.

Table 10. Individual States' Standardized Loadings on Correlated Components, Soybeans, 1975 to 1995 Period

State	Correlated Component		
	CC1	CC2	CC3
MISSOURI	0.106	0.216	-0.017
IOWA	0.089	0.229	0.063
ILLINOIS	0.103	0.136	0.022
MINNESOTA	0.079	0.187	0.114
OTHER STATES	0.090	-0.111	0.353
OHIO	0.083	0.119	0.028
TENNESSEE	0.096	-0.183	0.324
MISSISSIPPI	0.092	-0.191	0.221
NEBRASKA	0.085	0.084	-0.120
LOUISIANA	0.054	-0.057	0.209
INDIANA	0.087	-0.030	-0.087
SOUTH DAKOTA	0.049	0.095	-0.099
KANSAS	0.061	0.033	-0.124
KENTUCKY	0.101	-0.289	0.024
NORTH CAROLINA	0.077	-0.138	-0.044
ARKANSAS	0.078	-0.183	-0.052
WISCONSIN	0.048	-0.037	-0.072
NORTH DAKOTA	0.004	0.079	-0.006
MICHIGAN	0.057	-0.143	-0.113

U.S. Soybeans: Latter Time Period (1996-2017)

The four-fold cross-validation procedure produced an optimal CV-R^2 at the maximum of eight correlated components for the 1996 to 2017 period. The non-standardized and standardized coefficient values along with the indicators for level of significance and standardized share are shown in Table 11. The direct effects component (CC_1) has almost 60% of the total standardized share with the first three indirect effects components (CC_2, CC_3, and CC_4 each having around a 10% share. All of the coefficients are statistically significant at the 95% level or better with the exception of the final component (CC_8) which has a p-value (0.107) just outside the 90% confidence range.

Table 11. Correlated Component Regression and Standardized Coefficients, Soybeans, 1996 to 2017 Period

Correlated Component	Value	Standard Error	T-Statistic	Pr > |t|	Signif*	Standardized Value	Share of Standardized Total (%)
CC_1	0.131	0.001	119.999	< 0.0001	***	0.912	59.7%
CC_2	1.115	0.056	19.779	< 0.0001	***	0.152	9.9%
CC_3	0.726	0.060	12.026	< 0.0001	***	0.167	10.9%
CC_4	0.904	0.084	10.807	< 0.0001	***	0.157	10.3%
CC_5	0.403	0.061	6.548	< 0.0001	***	0.053	3.5%
CC_6	0.978	0.258	3.793	0.002	***	0.038	2.5%
CC_7	1.041	0.378	2.751	0.017	**	0.033	2.1%
CC_8	0.396	0.229	1.730	0.107		0.016	1.1%

*Three-stars = significantly different from zero at 99% confidence level, two-stars = 95%, one-star = 90%.

The component values over time are shown in Figure 6 and Figure 7. For the direct effects component (CC_1), the component has statistically significant and positive Spearman rank-order correlations with yield versus trend, annual average ONI, and 1st half ONI. It has negative correlations with extreme maximum temperature, days with minimum temperature under 32 degrees, days with maximum temperature over 90 degrees, and the extreme minimum temperature.

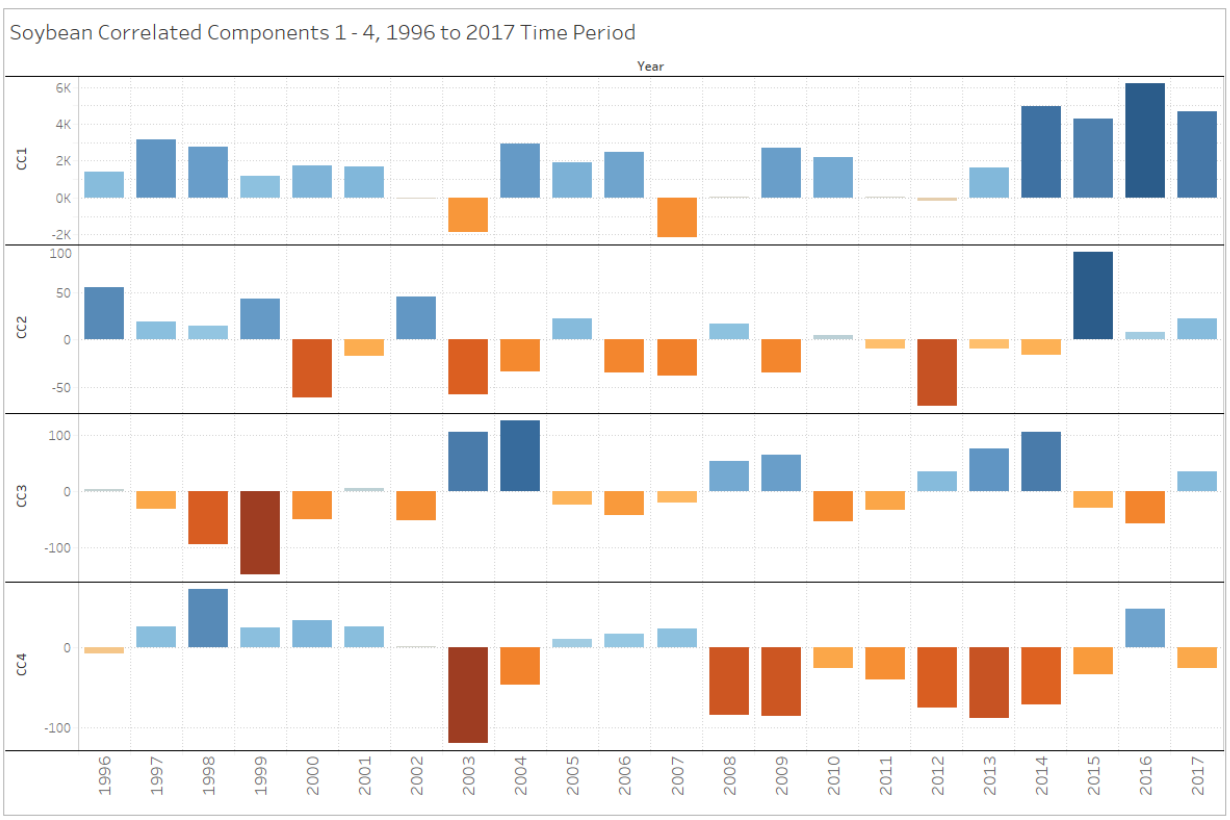

Figure 6. First Four Correlated Component Values for Soybeans, 1996 to 2017 Period

The first indirect effects coefficient (CC_2) does not have any correlation coefficients that are statistically significant (p-values less than 0.1); however, the most significant correlation is positive with the yield versus trend. This is followed by negative correlations with days with maximum temperature over 70 degrees, and the extreme minimum temperature. This component appears to be a reinforcement of the direct effects although note that it provides a very strong suppressive effect on the drought year of 2012 when compared to the direct effects component.

The second indirect effects component (CC_3) has significant negative correlations on average minimum temperature, average temperature, excess abandonment, and average maximum temperature. It has positive correlations on 2nd half ONI, heating degree days, and change in annual ONI. This component appears to have a suppressive effect related to weather variability particularly in terms of temperatures.

The third indirect effects component (CC_4) has significant positive correlation with average minimum temperature, average maximum temperature, and average temperature. It has significant negative correlation with the extreme maximum temperature and total snowfall in inches. Again, this component reflects weather variability and appears to offset many of the years for CC_3. So, it is likely the CC_3 and CC_4 reflect a combined effect related to weather extremes.

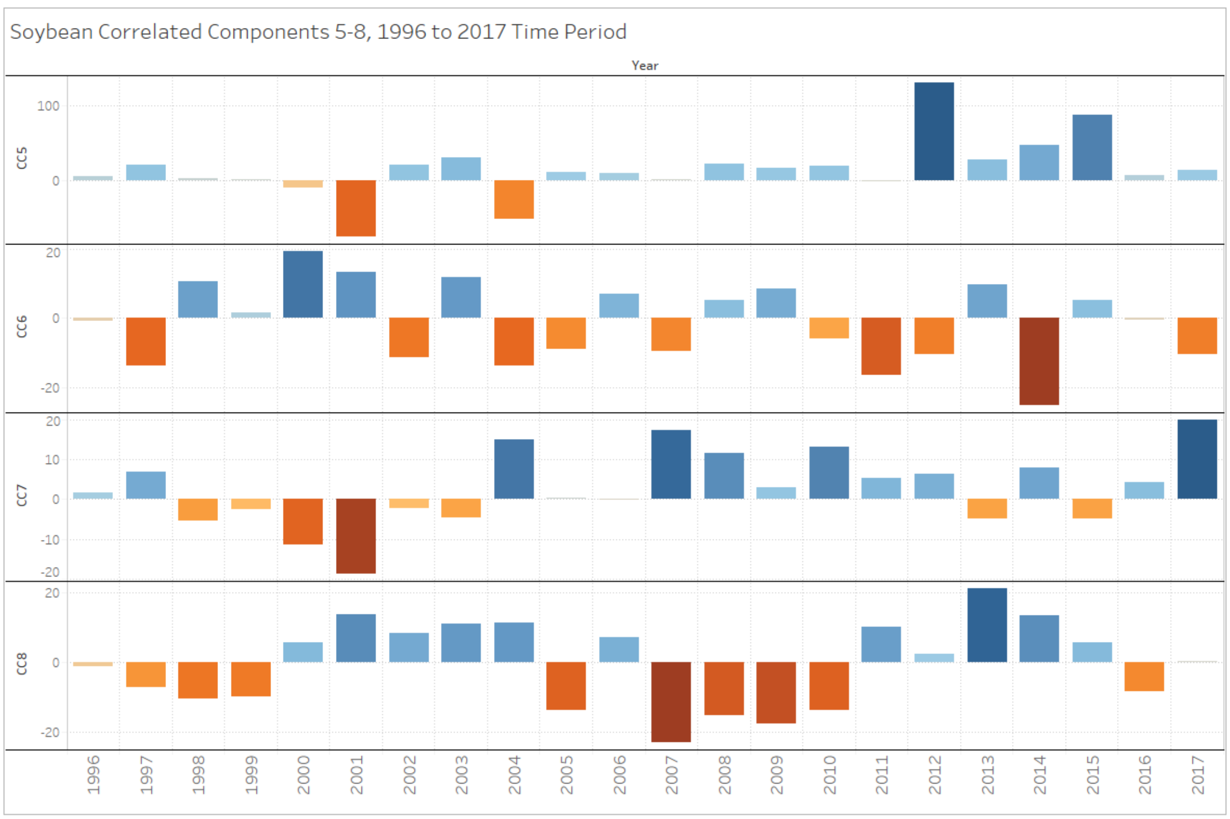

Figure 7. Last Four Correlated Component Values for Soybeans, 1996 to 2017 Period

The fourth indirect effects component (CC_5) has significant positive correlations with the planted acreage differential, 2nd half ONI, the extreme maximum temperature, and average annual ONI.

It also is negatively correlated with the extreme maximum snowfall. This component likely has effects tied to planting progress and early season growing conditions.

The fifth indirect effects component (CC_6) has only one significant correlation that is positively related to the extreme maximum snowfall. It also has negative correlations with extreme maximum daily precipitation and the planted acreage differential. As with CC_3 and CC_4, it appears to have an opposite relationship with CC_5 and is likely that CC_5 and CC_6 represent a combined effect.

The sixth indirect effects component (CC_7) has a significant positive correlation with the extreme maximum precipitation and total precipitation variables. Therefore, it is likely tied to the level of precipitation. The final indirect effects component (CC_8) has significant negative correlation with days with precipitation over 1.0 inches and total precipitation. It has a positive correlation with 2nd half ONI. As with the previous two pairs, it appears that CC_7 and CC_8 work together to reflect overall precipitation during the growing season.

Table 12. Individual States' Coefficient Values and Shares, Soybeans, 1996 to 2017 Period

Rank*	State	Coefficient	Standard Error	T-Statistic	Pr > |t|	Signif***	Standardized Coefficient	Share(%)*
1	IOWA	1.361	0.075	18.110	< 0.0001	***	0.259	15.9%
2	MINNESOTA	1.572	0.099	15.809	< 0.0001	***	0.204	12.5%
3	ILLINOIS	1.090	0.039	28.071	< 0.0001	***	0.200	12.3%
4	INDIANA	1.092	0.080	13.688	< 0.0001	***	0.102	6.3%
5	KANSAS	1.038	0.061	17.014	< 0.0001	***	0.094	5.8%
6	OHIO	1.264	0.072	17.480	< 0.0001	***	0.092	5.7%
7	TENNESSEE	1.903	0.205	9.290	< 0.0001	***	0.079	4.9%
8	MISSOURI	0.736	0.077	9.512	< 0.0001	***	0.078	4.8%
9	NEBRASKA	0.967	0.066	14.691	< 0.0001	***	0.077	4.7%
10	SOUTH DAKOTA	0.811	0.032	25.691	< 0.0001	***	0.074	4.6%
11	ARKANSAS	1.490	0.223	6.691	< 0.0001	***	0.072	4.4%
12	NORTH DAKOTA	1.009	0.036	27.797	< 0.0001	***	0.072	4.4%
13	MISSISSIPPI	1.685	0.162	10.401	< 0.0001	***	0.069	4.3%
14	WISCONSIN	-1.560	0.214	-7.279	< 0.0001	***	-0.059	3.6%
15	KENTUCKY	0.916	0.126	7.291	< 0.0001	***	0.039	2.4%
16	NORTH CAROLINA	0.915	0.151	6.072	< 0.0001	***	0.026	1.6%
17	OTHER STATES	0.230	0.112	2.055	0.0605	*	0.014	0.8%
18	LOUISIANA	-0.259	0.202	-1.282	0.2224		-0.009	0.5%
19	MICHIGAN	-0.182	0.150	-1.219	0.2446		-0.006	0.4%
	[Constant]	10.460	3.401	3.076	0.0088	***		
Herfindahl-Hirschman Index**								**838**

*Rank and share based upon absolute value of standardized coefficient.

**Equals sum of squared percentage shares (range from 0 to 10,000).

***Three-stars = significantly different from zero at 99% confidence level, two-stars = 95%, one-star = 90%.

Table 12 shows the individual states' non-standardized and standardized coefficients along with indicators for the level of statistical significance and standardized share. All of the states' regression coefficients are significant at the 99% level with the exception of the Other States (90% significant), and Louisiana and Michigan (not significant). Wisconsin is notable in having a highly significant but negative regression coefficient. In terms of the share of total standardized coefficient values, Iowa stands out in the top spot with 15.9% followed by Minnesota (12.5%) and Illinois (12.3%). There is a big gap before reaching the next tier of states (Indiana, Kansas, and Ohio) beginning at 6.3 percent. Note that the Herfindahl-Hirschman Index value of 838 represents an increase from the earlier 1975-1995 period.

Table 13. Individual States' Standardized Loadings on Correlated Components, Soybeans, 1996 to 2017 Period

State	Correlated Component							
	CC1	CC2	CC3	CC4	CC5	CC6	CC7	CC8
IOWA	0.120	0.771	-0.047	0.549	-0.048	-1.002	-0.109	-0.148
MINNESOTA	0.089	0.233	-0.128	0.241	0.610	0.997	0.222	-0.433
ILLINOIS	0.131	0.418	-0.044	-0.027	-0.013	0.373	0.260	0.423
INDIANA	0.121	-0.222	0.132	0.183	-0.353	-0.045	-0.359	0.414
KANSAS	0.100	0.147	0.026	-0.073	-0.152	0.050	0.080	-0.533
OHIO	0.093	-0.368	0.090	0.232	0.043	0.047	0.186	0.103
TENNESSEE	0.074	0.140	0.112	-0.221	0.157	0.341	-0.540	0.219
MISSOURI	0.105	-0.306	0.123	0.070	-0.310	-0.029	0.564	-0.253
NEBRASKA	0.097	0.159	-0.090	-0.054	-0.294	0.283	-0.251	0.061
SOUTH DAKOTA	0.105	0.051	-0.114	-0.060	0.060	0.044	-0.151	-0.080
ARKANSAS	0.026	0.197	0.127	-0.207	0.133	-0.225	0.740	0.445
NORTH DAKOTA	0.079	-0.158	0.038	0.084	0.159	-0.094	-0.043	0.054
MISSISSIPPI	0.034	0.222	0.125	-0.104	0.187	-0.360	0.278	-0.337
WISCONSIN	0.088	-0.149	-0.272	-0.314	-0.002	-0.529	-0.067	0.038
KENTUCKY	0.071	-0.094	0.086	-0.136	0.054	0.019	-0.360	0.207
NORTH CAROLINA	0.047	-0.258	0.139	0.017	0.125	0.001	-0.235	-0.160
OTHER STATES	0.038	-0.201	0.091	-0.219	0.107	0.383	0.317	-0.104
LOUISIANA	0.045	-0.083	0.055	-0.161	0.086	-0.612	-0.109	0.063
MICHIGAN	0.077	-0.224	-0.206	-0.115	0.112	-0.127	0.246	0.028

The individual states' standardized loadings upon the correlated components are shown in Table 13. Note that Wisconsin's negative influence can be traced to negative loadings on six out of the eight components with a strong negative loading on the weather variability (CC_3 and CC_4) and planting progress (CC_6) components.

U.S. Soybeans: Summary of Changes in State Rankings and Concentration

Table 14 shows a summary of the change in rankings and standardized shares by state along with the change in the Herfindahl-Hirschman Index. The biggest upward jumps in the rankings were Kansas (13th to 5th), Indiana (11th to 4th), North Dakota (18th to 12th), and Arkansas (16th to 11th). The largest downward jumps in the rankings were Other States (5th to 17th) and Louisiana (10th to 18th). In terms of standardized shares, Iowa (+6.3%) had the largest increase while the Other States (-7.1%) had the largest decrease. Note that the top state from the pre-GMO rankings, Missouri, saw a significant decline in both ranking (1st to 8th) and share (-5.0%).

Unlike corn, the HHI for soybeans increased by 158 points between the two periods which indicates an increase in concentration rather than a decline; however, this is a relatively minor increase by HHI standards. While this is somewhat surprising given the results for corn, it is likely due to the technological gains in the major producing states (Iowa, Illinois, and Minnesota) more than outweighing the distributive impacts of shorter maturing varieties and longer growing seasons in the northern and western states. Given the changes in rankings, as with corn, there appears to be a general shift from East to West in geographic importance (with the exception of Indiana).

Table 14. Change in State-Level Rankings and Shares, Soybeans

State	Rank			Share		
	1975 to 1995	1996 to 2017	Change	1975 to 1995	1996 to 2017	Change
MISSOURI	1	8	-7	9.8%	4.8%	-5.0%
IOWA	2	1	+1	9.6%	15.9%	+6.3%
ILLINOIS	3	3	0	8.9%	12.3%	+3.4%
MINNESOTA	4	2	+2	8.9%	12.5%	+3.7%
OTHER STATES	5	17	-12	7.9%	0.8%	-7.1%
OHIO	6	6	0	7.4%	5.7%	-1.8%
TENNESSEE	7	7	0	7.1%	4.9%	-2.2%
MISSISSIPPI	8	13	-5	5.8%	4.3%	-1.5%
NEBRASKA	9	9	0	5.6%	4.7%	-0.9%
LOUISIANA	10	18	-8	4.9%	0.5%	-4.3%
INDIANA	11	4	+7	4.6%	6.3%	+1.7%
SOUTH DAKOTA	12	10	+2	3.6%	4.6%	+1.0%
KANSAS	13	5	+8	3.3%	5.8%	+2.4%
KENTUCKY	14	15	-1	3.2%	2.4%	-0.8%
NORTH CAROLINA	15	16	-1	2.9%	1.6%	-1.3%
ARKANSAS	16	11	+5	2.3%	4.4%	+2.1%
WISCONSIN	17	14	+3	2.1%	3.6%	+1.6%
NORTH DAKOTA	18	12	+6	1.3%	4.4%	+3.2%
MICHIGAN	19	19	0	0.8%	0.4%	-0.5%
Herfindahl-Hirschman Index	**680**	**838**	**158**			

SUMMARY AND CONCLUSIONS

Over the past 20 years, U.S. agriculture has witnessed profound changes with respect to technology, climate, farm policy, and other factors (ethanol production, Chinese demand, etc.) that have major repercussions with regards to the geographic distribution of crop production. There have been many recent studies that have examined both the direct and indirect impacts of these production factors upon crop yields, acreage, and production from both a temporal and spatial perspective. However, little to no attention has been paid to the impact of these factors upon the relative influence of each individual state's crop production outcomes as they relate to the national outcome. This question has particular importance to those engaged in a wide range of crop production and marketing activities ranging from logistics to storage to research and development on crop technology and cropping alternatives. The question is also important from a risk management perspective since history has shown that major crop production events, both negative and positive, tend to be limited in geographical perspective. Therefore, the more concentrated the geographic importance of a particular crop's production, from a national perspective, the more significant the impact of any production event (drought, hurricane, etc.) if the event area contains the geographic locus of crop production from an importance perspective.

The purpose of this study is to address this question of state-level geographic importance for U.S. corn and soybeans by employing the following procedure. The regression results for corn indicated that Iowa, Illinois, and Minnesota were the primary states influencing corn production in the earlier 1975-1995 period with a 47.2% share of the standardized coefficient values. For the later 1996-2017 time period, Illinois, Nebraska, and Iowa occupied the three spots with a 37.8 percent share. Minnesota fell 10 places from 3rd down to 13th while Nebraska increased 14 places from 16th to 2nd. Other states making significant gains in the rankings included Texas, which rose 9 places from 19th to 10th overall. The HHI declined by 286 points (1103 to 818), which indicated a moderate increase in the dispersion of geographic influence from the former (1975-1995) to the latter (1996-2017) periods.

The regression results for soybeans indicated that Missouri, Iowa, and Illinois were the primary states influencing national soybean production in the 1975-1995 period with a much lower (than corn) 28.3 percent share of the standardized coefficient values. The HHI value of 680 for soybeans was also considerably lower than corn (1103) indicating that soybeans were considerably more diversified from a geographic perspective in the earlier period. For the latter 1996-2017 period, the top three states were Iowa, Minnesota, and Illinois with a 40.7 percent share of the standardized values. The residual (”Other States”) value had the largest decline in ranking falling 12 places from 5th to 17th. Louisiana (-8) and Missouri (-7) also saw major declines in ranking while Kansas (+8), Indiana (+7), and North Dakota (+6) had the largest increases in rankings. The HHI for the 1996-2017 period increased to 838 which was slightly larger than corn (818) and indicated a slight increase in the concentration of geographic influence among the top states for soybean production.

Overall, the results showed a shifting geographic dynamic for both corn and soybeans with the emphasis shifting from east to west in general direction. This makes intuitive sense as many of the observed technological and climatic changes over the past several decades point towards corn and soybean varieties that require a shorter growing season, and the increase in the number of frost-free days in many of the states in the northern reaches of the U.S. Corn Belt region. Additionally, the greater utilization of irrigation in crop production has likely contributed to the westward expansion of both corn and soybean production — often at the expense of wheat and cotton production. The slight decline in the HHI for corn indicates that production influence is becoming slightly more diversified from a geographic perspective. For soybeans, the opposite effect has occurred with a slight increase in the HHI pointing towards greater influence from the key producing states of Iowa, Minnesota, and Illinois — likely the result of a shift from corn to soybean acres as all three states lost influence shares in corn production between the two time periods.

References Cited

Adams, G., P. Westhoff, B. Willott, and R.E. Young. 2001. "Do 'Decoupled' Payments Affect U.S. Crop Area? Preliminary Evidence from 1997-2000." *American Journal of Agricultural Economics* 83(5):1190–1195.

Alkerwi, A., C. Vernier, N. Sauvageot, G.E. Crichton, and M.F. Elias. 2015. "Demographic and socioeconomic disparity in nutrition: application of a novel Correlated Component Regression approach." *BMJ Open* 5(5):e006814–e006814.

Andrews, D.W. 1993. "Tests for Parameter Instability and Structural Change with Unknown Change Point." *Econometrica* 60:821–856.

Bair, E., T. Hastie, D. Paul, and R. Tibshirani. 2006. "Prediction by Supervised Principal Components." *Journal of the American Statistical Association* 101(473):119–137.

Bruening, G., and J.M. Lyons. 2000. "The case of the FLAVR SAVR tomato." *California Agriculture* 54(4):6–7.

Bunge, J. 2018. "A warming climate brings new crops to frigid zones." *The Wall Street Journal*. Available at: https://www.wsj.com/articles/a-warming-climate-brings-new-crops-to-frigid-zones-1543168786 [Accessed November 26, 2018].

Burke, M., and K. Emerick. 2016. "Adaptation to Climate Change: Evidence from US Agriculture." *American Economic Journal: Economic Policy* 8(3):106–140.

Cai, R., J.D. Mullen, J.C. Bergstrom, W.D. Shurley, and M.E. Wetzstein. 2013. "Using a Climate Index to Measure Crop Yield Response." *Journal of Agricultural and Applied Economics* 45(04):719–737.

Chen, S., X. Chen, and J. Xu. 2016. "Impacts of climate change on agriculture: Evidence from China." *Journal of Environmental Economics and Management* 76:105–124.

Deschenes, O., and M. Greenstone. 2007. "The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather." *The American Economic Review* 97(1):354–385.

Du, X., C.L. Yu, D.A. Hennessy, and R. Miao. 2015. "Geography of crop yield skewness." *Agricultural Economics* 46(4):463–473.

Ezekiel, M. 1941. *Methods of Correlation Analysis* 2nd ed. New York: John Wiley and Sons.

Fei, C.J., B.A. McCarl, and A.W. Thayer. 2017. "Estimating the Impacts of Climate Change and Potential Adaptation Strategies on Cereal Grains in the United States." *Frontiers in Ecology and Evolution* 5:1–12.

Garcia, P., S.E. Offutt, M. Pinar, and S.A. Changnon. 1987. “Corn Yield Behavior: Effects of Technological Advance and Weather Conditions.” *Journal of Climate and Applied Meteorology* 26:1092–1102.

Garver, M.S., and Z. Williams. 2018. “Improving the validity of theory testing in logistics research using correlated components regression.” *International Journal of Logistics Research and Applications* 21(4):363–377.

Haile, M.G., T. Wossen, K. Tesfaye, and J. von Braun. 2017. “Impact of Climate Change, Weather Extremes, and Price Risk on Global Food Supply.” *Economics of Disasters and Climate Change* 1(1):55–75.

Harvey, D.I., S.J. Leybourne, and P. Newbold. 1998. “Tests for Forecast Encompassing.” *Journal of Business and Economics Statistics* 16:254–259.

Hendricks, W.A., and J.C. Scholl. 1943. “The Joint Effects of Temperature and Rainfall on Corn Yield.” No. 74, North Carolina Agricultural Experiment Station.

Houck, J.P., and P.W. Gallagher. 1976. “The Price Responsiveness of U.S. Corn Yields.” *American Journal of Agricultural Economics* 58:731–734.

Houseman. 1942. “Methods of Computing a Regression of Yield on Weather.” No. 302, Iowa Agricultural Experiment Station.

Huffman, W.E., Y. Jin, and Z. Xu. 2018. “The economic impacts of technology and climate change: New evidence from U.S. corn yields.” *Agricultural Economics* 49(4):463–479.

IPCC. 2015. “Climate change 2014: Synthesis Report. Contribution of Working Groups I, II, and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.” Intergovernmental Panel on Climate Change.

James, C., and A.F. Krattiger. 1996. “Global Review of the Field Testing and Commercialization of Transgenic Plants: 1986 to 1995 The First Decade of Crop Biotechnology.” ISAAA Briefs No. 1, ISAAA.

Kaufmann, R.K., and S.E. Snell. 1997. “A Biophysical Model of Corn Yield: Integrating Climatic and Social Determinants.” *American Journal of Agricultural Economics* 79(1):178–190.

Kennedy, P. 1998. *A Guide to Econometrics* 4th ed. Cambridge, MA: The MIT Press.

Kuhn, M., and K. Johnson. 2013. *Applied Predictive Modeling*. New York: Springer.

Kukal, M.S., and S. Irmak. 2018. “Climate-Driven Crop Yield and Yield Variability and Climate Change Impacts on the U.S. Great Plains Agricultural Production.” *Scientific Reports* 8(1):1–18.

Li, Y., R. Miao, and M. Khanna. 2019. “Effects of Ethanol Plant Proximity and Crop Prices on Land-Use Change in the United States.” *American Journal of Agricultural Economics* 101(2):467–491.

Lobell, D.B., M.J. Roberts, W. Schlenker, N. Braun, B.B. Little, R.M. Rejesus, and G.L. Hammer. 2014. “Greater Sensitivity to Drought Accompanies Maize Yield Increase in the U.S. Midwest.” *Science* 344(6183):516–519.

Lobell, D.B., W. Schlenker, and J. Costa-Roberts. 2011. “Climate Trends and Global Crop Production Since 1980.” *Science* 333(6042):616–620.

Magidson, J. 2010. “Correlated Component Regression: A Prediction/Classification Methodology for Possibly Many Features.” In *Proceedings of the American Statistical Association*.

Magidson, J., and K. Wassmann. 2010. “The Role of Proxy Genes in Predictive Models: An Application to Early Detection of Prostate Cancer.” In *2010 JSM Proceedings of the American Statistical Association, Biometrics Section*. pp. 2739–2753.

Marshall, E., M. Aillery, S. Malcolm, and R. Williams. 2015. “Agricultural Production under Climate Change: The Potential Impacts of Shifting Regional Water Balances in the United States.” *American Journal of Agricultural Economics* 97(2):568–588.

Mendelsohn, R., W.D. Nordhaus, and D. Shaw. 1994. “The Impact of Global Warming on Agriculture: A Ricardian Analysis.” *The American Economic Review* 84(4):753–771.

Menz, K.M., and P. Pardey. 1983. “Technology and U.S. Corn Yields: Plateaus and Price Responsiveness.” *American Journal of Agricultural Economics* 65(3):558–562.

Miao, R., M. Khanna, and H. Huang. 2016. “Responsiveness of Crop Yield and Acreage to Prices and Climate.” *American Journal of Agricultural Economics* 98(1):191–211.

O’Donoghue, E.J., M.J. Roberts, and N. Key. 2009. “Did the Federal Crop Insurance Reform Act Alter Farm Enterprise Diversification?” *Journal of Agricultural Economics* 60(1):80–104.

Quant, R. 1960. “Tests of the Hypothesis that a Linear Regression Obeys Two Different Regimes.” *Journal of the American Statistical Association* 55:324–330.

Rosenzweig, C., and M.L. Parry. 1994. “Potential impact of climate change on world food supply.” *Nature* 367:133–138.

Schlenker, W., W.M. Hanemann, and A.C. Fisher. 2006. “The Impact of Global Warming on U.S. Agriculture: An Econometric Analysis of Optimal Growing Conditions.” *The Review of Economics and Statistics* 88(1):113–125.

Schlenker, W., and M.J. Roberts. 2009. “Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change.” *Proceedings of the National Academy of Sciences* 106(37):15594–15598.

Shakya, S., W.W. Wilson, and B. Dahl. 2013. “Valuing New Random Genetically Modified (GM) Traits in Corn.” *Journal of Agricultural and Resource Economics* 38(1):107–123.

Tannura, M., S. Irwin, and D. Good. 2008. “Are Corn Trend Yields Increasing at a Faster Rate?” Marketing and Outlook Briefs No. MOBR 08-02, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign.

Tannura, M.A., S.H. Irwin, and D.L. Good. 2008. “Weather, Technology, and Corn and Soybean Yields in the U.S. Corn Belt.” Marketing and Outlook Research Report No. 1, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign. Available at: http://www.ssrn.com/abstract=1147803 [Accessed October 5, 2018].

Thompson, L.M. 1986. “Climatic Change, Weather Variability, and Corn Production.” *Agronomy Journal* 78(4):649–653.

Thompson, L.M. 1988. “Effects of Changes in Climate and Weather Variability on the Yields of Corn and Soybeans.” *Journal of Production Agriculture* 1(1):20–27.

Thompson, L.M. 1963. “Weather and technology in the production of corn and soybeans.” CARD Reports and Working Papers No. 17, Iowa State University, Center for Agricultural and Economic Development. Available at: http://lib.dr.iastate.edu/card_reports/17.

Thompson, L.M. 1969. “Weather and Technology in the Production of Corn in the U. S. Corn Belt.” *Agronomy Journal* 61(3):453–456.

Thompson, L.M. 1970. “Weather and Technology in the Production of Soybeans in the Central United States.” *Agronomy Journal* 62(2):232–236.

Tolhurst, T.N., and A.P. Ker. 2015. “On Technological Change in Crop Yields.” *American Journal of Agricultural Economics* 97(1):137–158.

Trivedi, J., and R. Birau. 2013. “Co-movements between emerging and developed stock markets in terms of global financial crisis.” In *1st WSEAS International Conference on Mathematics, Statistics, and Computer Engineering*. 1st WSEAS International Conference on Mathematics, Statistics, and Computer Engineering. Dubrovnik, Coatia: WSEAS, pp. 146–151.

Wallace, H.A. 1920. “Mathematical Inquiry into the Effect of Weather on Corn Yield in the Eight Corn Belt States.” *Monthly Weather Review* 48:439–456.

Wallander, S., R. Claassen, and C. Nickerson. 2011. “The Ethanol Decade: An Expansion of U.S. Corn Production, 2000-09.” Economic Information Bulletin No. 79, Economic

Research Service - USDA. Available at: http://www.ssrn.com/abstract=2131399 [Accessed October 5, 2018].