Fields from Afar: Evidence of Heterogeneity in United States Corn Rotational Response from Remote Sensing Data

Nicholas J. Pates* Nathan P. Hendricks†

Abstract

We construct estimates of own- and cross-price corn rotation elasticities using a field-level dataset that accounts for over 83% of the US corn-producing area. We allow rotational response to vary by estimating separate models across 115 subsamples that we delineate using Major Land Resource Areas (MLRAs) and soil characteristics. The results show a high degree of rotational response heterogeneity. Across the country, we find that rotational response is elastic in some areas and near zero in others. After aggregating the results to the national level, we find that modeling rotational response without allowing for heterogeneity produces a short-run own-price elasticity of corn planting of around 0.50 which conforms to the latest estimates in the literature. When allowing heterogeneous price sensitivity, our preferred estimate of the rotation elasticity is 0.69. This is evidence that imposing a uniform rotation response could seriously bias aggregate elasticity estimates.

Keywords: Rotation response, corn, heterogeneity
JEL codes: Q11, Q15, Q24

*Pates is an assistant professor in the Department of Agricultural Economics at the University of Kentucky. Department of Agricultural Economics, University of Kentucky, Lexington, KY 40503. npa286@uky.edu.

†Hendricks is an associate professor in the Department of Agricultural Economics at Kansas State University. Department of Agricultural Economics, Kansas State University, Manhattan, KS 66506. nph@ksu.edu.

Accurate estimates of supply elasticities are critical for studying a wide set of economic topics including the distribution of economic surplus, the allocation of production across producer groups, and the impact of production externalities brought about by price changes (Davis and Espinoza, 1998; Moschini, Lapan, and Kim, 2017; Wu et al., 2004). In this article, we estimate the rotational component of the supply response of corn across the United States and allow for heterogeneity in the relationship between planting decisions and prices. This helps correct for potential bias that can arise when rotational response is assumed to be uniform. Our model uses over 32 million observations from over 3.7 million fields within the US and accounts for 83% of corn production. We find a high degree of variation in rotational response across the US and that, on aggregate, this response is larger when weighted by field acreage than by field acreage and county yields. Our elasticity estimates are larger than those in the literature and we are the first to estimate a crop rotation elasticity at the national level using field-level data. We compare estimates from our heterogeneous-response model to estimates from a model that imposes homogeneous coefficients and find substantial downward bias of the aggregate planting elasticity in the homogeneous model.

Heterogeneity in rotational patterns and price response can bias models that combine observations together by pooling them into a single model. This bias arises in heterogeneous coefficient models when there is either heterogeneity in the effect of the lagged dependent variable or when there is autocorrelation in the other regressors (Pesaran and Smith, 1995). In our context of crop choices, heterogeneity in the dynamics arises from differences in producers' proclivity to rotate from one crop to another. This makes it difficult to isolate the influence of the lagged dependent variable from autocorrelated crop prices. Pesaran and Smith (1995) show that failing to incorporate parameter heterogeneity causes the estimate of the coefficient on the autocorrelated variable to be biased towards unity. Hendricks, Smith, and Sumner (2014) show how this pooling bias is accentuated when the data are aggregated to county-level panel data and gives the false impression that acreage responds gradually to price changes.

We make several important contributions beyond Hendricks, Smith, and Sumner (2014) and other recent supply response literature by considering heterogeneous acreage response to price across the entire country. First, we provide an improved estimate of national-level corn acreage elasticities by covering more than 80% of the nation's corn growing area instead of only Iowa, Illinois, and Indiana and by aggregating the heterogeneous acreage response using average corn yields and field sizes as weights. Our preferred estimate of the short-run (long-run) own-price elasticity is 0.69 (0.54) for the nation while Hendricks, Smith, and Sumner (2014) estimate an elasticity of 0.40 (0.29). The magnitude of these elasticities is critical for policy analysis. We also find that acreage response is larger in lower-yielding areas. This relationship could be important for understanding changes in aggregate crop productivity. Because it produces a disproportional share of the nation's corn, much of the corn supply response literature focuses its efforts on the Corn Belt and the state of Iowa (Lee and Helmberger, 1985; Tegene, Huffman, and Miranowski, 1988; Orazem and Miranowski, 1994; Langpap and Wu, 2011; Hendricks, Smith, and Sumner, 2014; Kim and Moschini, 2018). Our results emphasize the importance of modeling the entire growing area rather than using a region as representative.

Second, we find substantially more spatial heterogeneity in elasticities. We find similar acreage elasticities in the central Corn Belt as Hendricks, Smith, and Sumner (2014), but the national-level own-price elasticity is larger due to a more elastic response on the fringes of the Corn Belt. The spatial heterogeneity in planting response is relevant for assessing the environmental impacts caused by policies or shocks that change prices. If there is spatial heterogeneity, then researchers need to account for the correlation between planting response and environmental sensitivity to accurately assess the environmental impact. Furthermore, heterogeneous rotational response implies that a national level supply response cannot necessarily be applied to study the local environmental impacts. While assessing environmental impacts from changes in corn prices is beyond the scope of this article, there is evidence that rotation response is not independent of the environmental sensitivity of the area. For example, corn acreage within the Prairie Pothole Region—a prominent breeding area for migratory waterfowl—is around three times more sensitive to corn price changes than in the central Corn Belt.

Third, we find that ignoring parameter heterogeneity results in substantial pooling bias. Pesaran and Smith (1995) note that pooling bias increases with the degree of parameter heterogeneity so utilizing a country-level rather than region-level analysis better illustrates the importance of modeling heterogeneity to reduce bias. Using a pooled model, we estimate a short-run (long-run) own-price elasticity of 0.50 (0.42)—biased downwards by roughly 16 to 19 percentage points.

The recent supply response literature places more emphasis on incorporating heterogeneity (Lacroix and Thomas, 2011; Motamed, McPhail, and Williams, 2016; Haile, Kalkuhl, and von Braun, 2016). Many do this using fixed effects or additive separable effect frameworks. Models allowing for heterogeneous supply response coefficients are also becoming more popular as more evidence comes to light (Koutchadé, Carpentier, and Femenia, 2018). Seen as a major component of supply response, more of the literature also incorporates rotational frameworks (Hendricks, Smith, and Sumner, 2014; Hendricks et al., 2014; Langpap and Wu, 2011; Claassen, Langpap, and Wu, 2017). Variants of multinomial discrete choice models are popular ways to include the influence of rotations while allowing for heterogeneity.

There are practical trade-offs between modeling specific rotation practices and modeling across larger and more diverse areas. Random parameter and latent class multinomial discrete choice models are popular since they can consider many crop alternatives, allow for heterogeneous response, and model sequential cropping decisions (Langpap and Wu, 2011; Claassen, Langpap, and Wu, 2017). The complexity of these models and, for the mixed logit, their reliance on simulation, creates problems when estimating heterogeneous supply response in very large datasets. While highly flexible, using models like the mixed logit and the nested logit requires estimating more parameters and, as a result, are more difficult to scale to larger field-level datasets. Many previous studies that used the nested or mixed logit models had fewer than 100,000 observations (Koutchadé, Carpentier, and Femenia, 2018; Langpap and Wu, 2011; Claassen, Langpap, and Wu, 2017).

For this study, we adopt the framework of Hendricks, Smith, and Sumner (2014) and Hendricks et al. (2014). They use a Markov-chain framework and model rotations with two or more individual discrete choice models. Each of these models characterize the probability of a different crop transition. They allow for heterogeneity by estimating separate models using different subsets of the dataset. We separately estimate models by MLRAs and soil characteristics. Since MLRAs differ by climate, geography, and physiography we allow variable response across agriculturally relevant dimensions. While this method requires more regressions, each regression is computationally simple and better lends itself to parallelization. This approach can estimate heterogeneous rotational response while maintaining tractability on large datasets (Hendricks, Smith, and Sumner, 2014; Hendricks et al., 2014).

We are able to deliver these results with the use of remote sensing satellite data. Donaldson and Storeygard (2016) provide an in-depth discussion on the promise and challenges of remote sensing datasets to economic analysis. These datasets provide a geographically rich and temporally consistent set of observations across the wide expanse of the United States for over 9 years in the case of this study. This enables us to address relevant issues of pooling bias of dynamic crop rotation models that would not be possible otherwise. It also allows us to more precisely control for factors at the field-level as opposed to approximating them with aggregate measures. With these benefits come challenges. As these data are remotely observed, they are more prone to measurement and classification error than verified, ground-truthed data. In addition, due to the possibility of spatial correlation, it is less clear when one can consider one observation as independent of another. Our research design attempts to address these issues by dividing our data across theoretically important spatial dimensions and accounting for possible within-year correlation when estimating standard errors. These data classified corn plantings relatively well but were less accurate in classifying specific corn alternative such as wheat and fallow.[1] We address potential classification error problems by focusing our research on the response of corn only and grouping corn alternatives together into a single "other" category. Finally, we address observational dependence by modeling at the field-level as defined by Farm Service Agency records.

Conceptual Framework

We can decompose the crop supply function using a simple product of parcel size, cropping probabilities, and the crop yield per unit area of the parcel shown in equation 1. Where $\mathbb{P}_{ik}$ is the probability that crop k is selected given parcel i is dedicated to cropping activities, $\mathbb{P}_i^{crop}$ is the probability that parcel i is used for cropping activities, $acre_i$ is the acreage of parcel i, and $yield_{ik}$ is the per-acre output for crop k on parcel i.

$$(1) \quad Q_{ik} = acre_i \mathbb{P}_{ik} \mathbb{P}_i^{crop} yield_{ik}$$

The effect of a price change on the supply of crop k is shown in equation 2. Under this characterization, supply response is made up of a rotational response, an extensification response, and a yield response shown as the three respective terms in the equation.

$$(2) \quad \frac{\partial Q_{ik}}{\partial P_j} = acre_i \left[\underbrace{\frac{\partial \mathbb{P}_{ik}}{\partial P_j} \mathbb{P}_i^{crop} yield_{ik}}_{\text{Rotational Response}} + \underbrace{\mathbb{P}_{ik} \frac{\partial \mathbb{P}_i^{crop}}{\partial P_j} yield_{ik}}_{\text{Extensification Response}} + \underbrace{\mathbb{P}_{ik} \mathbb{P}_i^{crop} \frac{\partial yield_{ik}}{\partial P_j}}_{\text{Yield Response}} \right]$$

The rotational response refers to how often crop k is planted within a sequence of crop choices. The extensification response refers to changes in the total area engaged in cropping activities. While the physical sizes of parcels do not change, the proportion of parcels in cropland can change. Lastly, the yield response is the change to the per-acre output from price changes due to, for example, changes in inputs applied.

We focus exclusively on the rotational response, the frequency that a given crop is planted within a sequence. Though it is a partial response model, the literature suggests that the planting response component comprises the majority of supply's overall response to price changes. Because converting land from non-cropland to cropland is a slow and expensive process, transitory price changes are likely to have small impacts on total corn area through cropland expansions. Barr et al. (2011) found that the extensive land change was quite inelastic to price changes. Miao, Khanna, and Huang (2016) find that yield response constitutes around 30% of overall supply response. Others find minimal yield response to price (Berry and Schlenker, 2011).

Equation 3 shows the expected supply response arising from intensive planting changes, since the probabilities and yields are random variables.[2]

$$\mathbb{E}\left[\frac{\partial Q_{ik}}{\partial P_j}\right] = \left(\mathbb{E}\left[\frac{\partial \mathbb{P}_{ik}}{\partial P_j}\right]\mathbb{E}\left[\mathbb{P}_i^{crop} yield_{ik}\right] + Cov\left(\frac{\partial \mathbb{P}_{ik}}{\partial P_j}, \mathbb{P}_i^{crop} yield_{ik}\right)\right) acre_i \tag{3}$$

Equation 3 helps illustrate the error that ignoring response heterogeneity introduces. Using a single rotational estimate, such as an average of the responses, for all parcels is only valid if the price effect on the conditional probability of planting corn is uncorrelated with the product of yields and the cropping probability.

Data

In this section we describe the sources and characteristics of the datasets we use in our empirical models. The data in this study serve three purposes: (1) identifying crop choices of individual producers over time, (2) incorporating relevant heterogeneity across the United States, and (3) providing independent price variables and regression controls.

Crop Classification

We identify land cover using the USDA's Cropland Data Layer (CDL). This dataset identifies land cover at the sub-field level from as far back as 1997 to present. The CDL classifies land cover using satellite imagery, which provides a categorical raster image of the US at a 30-meter resolution. While some CDL observations date back to the 1990s, local crop price data restrict the dataset between 2004 and 2016.[3] Linking the CDL observations to the 2008 Common Land Unit (CLU) shapefile dataset provides field-level crop-choice observations across the United States (Woodard, 2016). The USDA delineates CLUs as the smallest contiguous land pieces that are associated with USDA farm programs with a common owner and a common producer. We utilize the 2008 version of the CLU map since this is the only year of data available to us. We fill in areas of the country with missing common land unit boundaries with polygons from Yan and Roy (2016). We merge the CDL with fields by identifying the crop at a point within the field boundary.

To incorporate heterogeneity, we estimate separate models over a set of geographic boundaries known as Major Land Resource Areas (MLRAs) established by the Natural Resource Conservation Service (NRCS). The set of MLRAs consists of 278 subregions within the US, delineated using agriculturally relevant features including physiography, geology, climate, water, soils, biological resources, and historic land use. This provides a convenient way of establishing meaningful heterogeneity in row crop agriculture across the country.

We use soil texture statistics to incorporate model heterogeneity within each MLRA. Specifically, we use the 5-group soil taxonomy classification provided by Benham, Ahrens, and Nettleton (2009). While the 12-class soil taxonomy is more commonly known, the 5-class aggregation provides enough observations in each class to robustly estimate separate regressions. We pool observations by soil class in instances where the within-MLRA texture classification groups training samples were too small to reliably estimate the models. If a soil texture group within an MLRA had less than 20,000 observations, we combined it into the next "closest" soil texture group in the MLRA.[4]

Due to the diversity of suitable crops, constructing a parsimonious rotational response model over an area as large as the United States is challenging. The CDL identifies a variety of crops but crop prices are more difficult to find. Many crops do not have associated futures contracts. This makes expected prices difficult to construct. Since we account for heterogeneity using pooled regressions at the MLRA-level, we need to ensure each regression has enough observations to produce reliable estimates. The Markov transition probability regression modeling strategy requires that the sample data consists of only observations with two consecutive row crop choices which can constrain the MLRA sample size.

We divided the crop observations into 5 groups: corn, priced corn alternative (PCA) crops, non-priced corn alternative (NPCA) crops, remaining crops (RCs), and non-cropland.[5] Corn consists of observations which the CDL classifies as conventional corn. That is, "corn" does not include double-cropping observations involving corn (e.g. double-cropping corn and soybeans) or less conventional varieties such as sweet corn or popcorn. Since our goal is to understand conventional corn response, in our empirical models we consider PCAs and NPCAs as "other" crops and corn as its own separate category. Priced corn alternatives (PCA) are CDL classified crops for which we have a measure of expected prices. The PCA prices enter the "other" crop price index value (soybeans, rice, non-Durham wheat varieties, and cotton) in our empirical models. PCAs also include associated double-cropped observations with these crops (e.g., winter wheat-cotton double-cropped observations). Non-priced corn alternatives (NPCA) crops are substitutes in production to corn and the PCAs that we do not have a measure of the expected price. NPCAs include small grain crops, oilseed crops, root crops, and fallow. The remaining crops (RC) category contains crops that are less substitutable to corn. This category includes specialized fruit and vegetable crops and perennial crops such as alfalfa. The final category, non-cropland, contains land uses that are not immediately suitable for crop production including pasture, forests, marshland, and developed lands.

To ensure we have adequate data to properly estimate the corn rotational response in each area, we use these classifications to remove MLRAs using a set of three hurdles. The first hurdle filters out MLRAs with less than 20% of its total acreage in corn, PCAs, NPCAs, or RCs to remove areas with low agricultural activity such as desert, mountainous, or developed areas. The second hurdle ensures that the price index reasonably applies to relevant alternatives to corn production. Since we construct the price index using prices from the PCAs to represent the prices of alternative crops, the second hurdle removes MLRAs if less than 50% of its combined PCA and NPCA acreage consists of PCA acreage. Last, since corn is the crop of interest, the third hurdle ensures the MLRAs have enough corn observations. This threshold removes MLRAs if corn makes up less than 10% of their combined corn, PCA and NPCA acres. In addition to these three hurdles, we also remove MLRAs with less than 50,000 total observations and MLRA groups where less than 20,000 observations enter either of the Markov transition regressions. We also exclude fields that are smaller than 15 acres because smaller fields are more prone to crop classification measurement error. After filtering, the data includes 68 MLRAs, 115 MLRA-soil texture groups and over 32 million individual observations.

Crop Prices

The importance of including expected prices and the debate on how to incorporate expectations has been a persistent issue in the agricultural supply response literature (Nerlove, 1956; Haile, Kalkuhl, and von Braun, 2016; Miao, Khanna, and Huang, 2016; Roberts and Schlenker, 2013; Gardner, 1976). A common theme is that when modeling planting decisions, it is important to use prices that reflect harvest-time expectations at or before the time at planting. Using the lagged harvest price was the earliest and simplest way of incorporating expectations in prices. In Nerlove's famous supply response article, he used an adaptive expectations model to produce expected prices (Nerlove, 1956). Others used futures prices since, under the efficient market hypothesis, these prices should reflect information about expected price changes (Gardner, 1976; Haile, Kalkuhl, and Braun, 2014).

We represent price expectations using a combination of the local spot price and the difference between the harvest and nearby futures prices trading before the corn-planting months. Roberts and Schlenker (2013) argue that the futures prices could be endogenous if futures prices include expected supply shocks. Roberts and Schlenker (2013) cite the competitive storage model and argue that past production could influence contemporaneous prices for storable commodities. They use lagged weather as an instrument for the current year's futures price. Another concern is that local prices could be endogenous if the acres of corn produced locally impacts the spot price. These potential sources of endogeneity are not likely to be a serious issue in our application for a few reasons. First, we model planting response at the disaggregated MLRA-soil-group level, allowing for distinct responses between groups. If there exists local endogenous drivers of price due to latent producer characteristics, then separately estimating planting response by MLRA and soil groups would mitigate this endogeneity. Second, after decomposing the variance of our prices, we find that 92% of the variance of corn prices and 90% of the other prices was temporal. This suggests that potential endogeneity stemming from local sources is not likely a practical concern. By including prices observed before planting decisions are made and including the basis effects from the nationally traded futures prices, we also help avoid possible temporally-related endogeneity issues. A final advantage that using futures price observations at the pre-plant stage of the season is that it allows us to better approximate what the market expects harvest-time prices to be at the time producers are developing their planting strategies. Hendricks, Janzen, and Smith (2014) found that modeling with futures prices trading before planting avoids most of the endogeneity concerns of futures price with growing area.

For much of the nation, corn planting takes place between early March and late April. We assume that the planning process begins in the months of January and February as this gives time for required crop-specific pre-plant land preparation and seed purchases. To construct expected prices, we first average local daily spot prices over the course of the months of January and February which we call the planting price $\left(C_{it}^{P}\right)$. Here the "P" signifies we are observing the statistic as the average value in the pre-plant months of January and February. Next, we average the daily nearby futures contract price and the harvest-time futures contract price for the respective commodities in January and February $\left(F_t^{PN}\text{ and }F_t^{PH}\text{ respectfully}\right)$. We construct the expected harvest-time spot prices according to equation 4.

$$(4)\quad E_{it}\left[P_{it}^{H}\right] = F_t^{PH} + \underbrace{\left[C_{it}^{P} - F_t^{PN}\right]}_{\text{Basis}} = \underbrace{\left[F_t^{PH} - F_t^{PN}\right]}_{\text{Expected Cost of Carry}} + C_{it}^{P}$$

We first compute the average annual nearby basis at planting and add the harvest-time futures price. The pre-plant average nearby basis incorporates the local basis pattern of the individual market, providing an estimate for basis at harvest time. Adding the expected basis to the expected harvest futures price gives the projected harvest-time local market spot price.

Another way to conceptualize the expected price is to note that the difference between the harvest-time contract price and the nearby contract price is the cost of carry for the commodity. This is the market's expected cost that a farmer would incur if she were to store grain until harvest time. Therefore, to incentivize a producer to deliver stored grain at the harvest date, she would require the current price plus the expected cost of storing the grain until harvest time. In efficient markets where the value of stored and "new" grain are the same, adding the cost of carry to the current cash price gives the expected price at harvest.

Our study considers price data for corn, soybeans, hard red winter wheat (HRWW), hard red spring wheat (HRSW), soft red winter wheat (SRWW), rice, and cotton. These prices are quoted in dollars per bushel with the exceptions of rice and cotton, which are in dollars per pound. Daily futures prices are available through the Data Transfer Network (DTN). The local spot price data are from the propriety datasets from the Data Transfer Network (DTN) and Cash Grain Bids (CGB) which we accessed through a Bloomberg terminal.[6]

To ensure that the average price estimates are representative of the pre-plant stage, we remove markets with less than 10 spot-price observations between January and February. From 2004 to 2016 there were 1,367 corn, 1,252 soybean, 84 HRSW, 96 HRWW, and 123 SRWW price locations. Coverage for the continuously observed local markets densely covers most of the major field crop production areas. We observe rice and cotton prices at the national level from 2004 to 2016. For cotton, we use daily observations of the USDA middling grade cotton average spot prices. For rice, we use average monthly observations from the National Agricultural Statistics Service's (NASS) Quickstats at the national level. Figure 1 shows the price coverage by crop.

With the center of the market city as a reference, we use our geo-located prices to construct annual basis maps for each commodity over the contiguous United States. After estimating the expected commodity price for each market, we interpolate these prices using ordinary kriging. We estimate basis map values as a raster with a resolution of 0.01 square degrees (or approximately 36 square miles). To maintain consistency in the price expectations estimates over time, the original observation set contains only markets with continuously observed price averages in every year from 2004 to 2016.

We simplify the study by considering a single crop, corn, examining rotations between corn and some other crop. As a reminder, these other crops consist of crops that are reasonable alternative choices to corn (PCAs and NPCAs). We characterize the "other" crop price using a weighted average of soybeans, HRWW, SRWW, HRSW, cotton, and rice prices. Using a common weighting scheme for every field is problematic since the set of relevant alternative crops to corn production differ across the country. For instance, cotton may be a relevant alternative crop in the Mississippi River Delta area but is irrelevant in Wisconsin. To account for this, we represent the other crop price as a Laspeyres index where the price weights differ by MLRA. Mathematically, this is the same index that underlies the Consumer Price Index and creates a "basket" of commodities indexed from $k = 1, \ldots, K$ using the quantities in some base period (period 0) to track the basket's changing prices. Equation 5 shows the functional form where p_{tk} is the price for commodity k at time t and q_{0k} is the total quantity of crop k produced in period 0.

$$(5) \quad P_t^O = \frac{\sum_{k=1}^{K} p_{tk} q_{0k}}{\sum_{k=1}^{K} p_{0k} q_{0k}}$$

Using unique price indices for each MLRA ensures that the other crop price consists of crops grown in the region. For instance, the dominant alternative crop in the state of Iowa is soybeans. If soybeans are the only alternative crop in a region, then $q_{k'} = 0 \ \forall k'$ where k' consists of non-soybean crops. This means that only soybean prices would enter the price index.

There are several complications to using the standard Laspeyres index in this study. First, crop choices are subject to change over time so it is unclear whether using the observed quantity produced in a single period q_0 represents the typical crop choice basket over the course of the study. Second, the analysis is at the field-level where we have only a single crop choice observation in each year. A particular crop planted at the beginning of the analysis does not preclude another crop from being considered in the future. Last, while we observe crop choices on some fields before 2008, the CDL dataset did not gain full coverage until 2008. To address these issues, we define q_{0k} as the MLRA-specific total production of crop k from 2008 to 2016. We compute total production by each MLRA for each crop by merging the field-level data with a county-level yield dataset, multiplying the yield with the field-level acreage choices, then summing over each MLRA from 2008 to 2016. Average annual yield data for each of these crops are available at the county-level through the National Agricultural Statistics Service (NASS).[7]

Field-Level Controls

Finally, we use field-level controls to incorporate individual field heterogeneity. To further control for soil differences at the field-level, we include the National Commodity Crop Productivity Index (NCCPI) and field slope as regressors. The NCCPI index takes many facets of the soil's crop productivity into account (Dobos, Sinclair, and Hipple, 2008). We also include the slope of the field since it is a key determinant of erosion and is a common control in the literature (Wang et al., 2015; Wu et al., 2004). We control for irrigation by including a 2012 Moderate Resolution Imaging Spectroradiometer Irrigated Agriculture Dataset (MIrAD) dataset which gives information on the irrigation status over the conterminous United States. This variable equals one if the field is irrigated and zero otherwise (Brown and Pervez, 2014). These data are derived from peak NDVI measures over cropland in each county in 2012, arranging the peak NDVIs, and selecting the irrigated peak NDVI threshold to calibrate irrigated area with the amount of land designated as irrigated in the 2012 Census of Agriculture (Pervez and Brown, 2010). There is potential that crop prices may be correlated with irrigation status. However, this is likely not a serious issue since we model separately within each MLRA and soil texture group. At the local scale we would expect that prices are spatially correlated according to distance to market as opposed to irrigation status. Another source of potential endogeneity could be if irrigation status is correlated with land quality. We mitigate this concern by including controls for land quality in our model. Our primary interest in this study is to estimate corn planting price elasticities and not the direct effect of irrigation, so we expect any remaining endogeneity concerns to have minimal impacts on our main findings.

Since we analyze planting decisions, it is important to control for pre-plant weather conditions. Extremely wet conditions can delay planting and cause farmers to plant alternative crops such as soybeans with later planting dates. Extremely dry conditions may cause farmers to shift to more drought-resistant crops. We incorporate extreme planting precipitation conditions by including two indicator variables into the models. The first indicator variable is a measure of exceptional dry planting conditions and equals one if the field's April-May precipitation was at or below the 25th percentile.[8] The second indicator variable indicates wet planting conditions and equals one if the field's April-May precipitation was above the

75th percentile. We use the annual historical data between 1983 and 2016 to define these percentiles. To account for possible agriculturally relevant interactions with weather, we construct these percentiles for each MLRA.

Data Summary

Table 1 shows summary statistics for our dataset. The data contain over 30 million observations across 115 different MLRA-soil texture groups across the country. Our total field acreage amounts to over 220 million acres which accounts for roughly 70% of cultivated cropland area in the US. The probability of planting corn was around 40% on average. This is a reasonable figure as corn was often rotated with at least one other crop each year and rotated less frequently in areas outside of the nation's Corn Belt. Our mean field size was around 60 acres. A little over 8% of the fields in the dataset were classified as irrigated and, on average, fields had an NCCPI value of around 0.6. Since the NCCPI measure is between 0 and 1 with larger values indicating better productivity, we can conclude that the average field had moderately high productivity potential. Our mean field slope of 3.1° suggests that the average field is gently sloping. Our extreme pre-plant precipitation indicators show that on average, 43% of our observations were in a relatively dry year and around 48% were in a relatively wet year. While we constructed these variables using the 25th and 75th precipitation percentiles, the percentiles relate to the past 30 years of weather observations. We can therefore interpret that precipitation in the pre-plant period was more volatile over the years in our study.

We also include a comparison of our observed corn acreage with official records by year. Because our transition data did not reach nationwide coverage until 2009, we compare our observed values between 2009 and 2016 with those from the National Agricultural Statistics Service (NASS) survey values in table 2. Our dataset accounts for roughly 83% of the total planted corn acreage in the country, consistently monitoring between around 70 and 80 million acres each year. The year-to-year consistency of our dataset's share of total acreage indicates that it is representative of major national corn planting trends.

Empirical Model

To estimate the heterogeneous response of rotations to price across the United States, we construct a set of Markov transition regression models over different subsets of the national sample. We estimate separate regressions by Major Land Resource Areas (MLRAs) and soil texture classifications group (g) to allow for distinct effects across the country. Following Hendricks et al. (2014), we estimate two Markov transition equations, use these results as elements of the transition matrix, and derive rotation probabilities. The first equation models the probability of planting corn given corn was previously cropped. The second equation estimates the probability of planting corn given some other crop was previously planted. The left-hand-side terms of the two equations are the conditional probabilities of planting corn and characterize the first-order Markov process. Here y_{it} equals one if corn is planted on field i at time t and zero otherwise. Using the estimated Markov process, we can estimate the unconditional probability of planting corn, the unconditional probabilities of planting a sequence of crops, and the probabilities of rotations involving corn and other crops.

Equations 6 and 7 show the structure of the Markov transition probabilities.

$$(6) \quad \phi_{it}^{CC} = Prob\left[y_{it} = 1 \mid y_{it-1} = 1\right] = \Lambda\left(\beta_{10} + \beta_1^C P_{it}^C + \beta_1^O P_{it}^O + \gamma_1 \mathbf{X}_{it}\right) \mid y_{it-1} = 1$$

$$(7) \quad \phi_{it}^{OC} = Prob\left[y_{it} = 1 \mid y_{it-1} = 0\right] = \Lambda\left(\beta_{20} + \beta_2^C P_{it}^C + \beta_2^O P_{it}^O + \gamma_2 \mathbf{X}_{it}\right) \mid y_{it-1} = 0$$

Here P_{it}^C and P_{it}^O are the expected harvest-time corn and other crop prices respectfully, and $\mathbf{X}_{it}$ contains field-level controls. These controls include static variables such as field

i's slope, soil productivity, and irrigation status and dynamic variables such as the yearly extreme pre-plant precipitation indicators and a time trend. Only fields growing either corn or other crops between the two consecutive periods are in the sample used to estimate these models. Here $\Lambda(\cdot)$ is the logistic function where $\Lambda(x) = \frac{1}{1+\exp\{-x\}}$. Equations 6 and 7 only differ by the data that enters each regression. Equation 6 uses only fields that planted corn in the previous period and equation 7 uses fields that had an other crop planted in the previous period.

Pooling observations together to estimate a single set of coefficients creates bias in panel regressions where there is lagged dependence like ours. Pesaran and Smith (1995) refer to this as "pooling bias." In the supplementary online appendix, we provide more details on the source of this bias by reframing our model as a single equation with a lagged dependent variable. Pooling bias arises when either there is unmodeled heterogeneity in the lagged dependent relationship or if the other regressors are autocorrelated. Ignoring heterogeneity in the lagged dependent relationship means that the coefficient on the lagged dependent variable effectively enters the error term and is correlated with the lagged dependence. If the model also has an autocorrelated regressor, then it will be impossible for the model to distinguish the effect of the autocorrelated regressors from the effect of the lagged dependent variable. The coefficient on the lagged dependent variable in the pooled model is biased upward and the coefficient on the autocorrelated regressor is biased downward. In our context, the incentives to rotate crops create a lagged dependence between contemporaneous and past crop choices and the key autocorrelated regressors are crop prices. To reduce pooling bias, we separately estimate equations 6 and 7 for each MLRA-soil texture group.

Equation 8 shows the structure of the Markov chain. In these equations, $(\mathbb{P}_{it})$ is the steady-state probability of planting corn.

$$(8)\quad \begin{bmatrix} \phi_{it}^{CC} & \phi_{it}^{OC} \\ \left(1-\phi_{it}^{CC}\right) & \left(1-\phi_{it}^{OC}\right) \end{bmatrix} \begin{bmatrix} \mathbb{P}_{it} \\ \left(1-\mathbb{P}_{it}\right) \end{bmatrix} = \begin{bmatrix} \mathbb{P}_{it} \\ \left(1-\mathbb{P}_{it}\right) \end{bmatrix}$$

As equations, 6 and 7 suggest, the left-hand-side variables in the transition matrix in equation 8 represent the probabilities of planting a given sequence conditional on the prior planting decision. For instance ϕ_{it}^{CC} is the probability of corn being planted given corn was planted in the prior year. Similarly, ϕ_{it}^{OC} is the probability that corn is planted conditional on an other crop being planted in the prior season.[9] The steady-state probability of planting corn in period t takes the form of equation 9.[10]

$$(9)\quad \mathbb{P}_{it} = \frac{\phi_{it}^{OC}}{1-\phi_{it}^{CC}+\phi_{it}^{OC}}$$

With the structure of the Markov chain, we define the short- and long-run effects of price changes on the probability of planting corn. To simplify notation, we drop the field and time subscripts. The state probability, shown in equation 10, is made up of two sequential probabilities. Some portion of those planting corn planted corn in the previous period. Therefore the state probability contains the probability of a corn-corn sequence $\mathbb{P}^{CC}$. The other portion of corn planters planted another crop in the previous period so the other-corn sequential probability $\left(\mathbb{P}^{OC}\right)$ is also in the state probability.

$$(10)\quad \mathbb{P} = \underbrace{\mathbb{P}\phi^{CC}}_{\mathbb{P}^{CC}} + \underbrace{\left(1-\mathbb{P}\right)\phi^{OC}}_{\mathbb{P}^{OC}}$$

Under short-run marginal effects, price changes only impact the state probability of planting corn through the transition matrix. Equation 11 shows the short-run marginal effect when the price of crop k changes. Here $\frac{\partial \phi^{CC}}{\partial P^k}$ is defined as the average partial effect of a crop-k price change on ϕ^{CC} and $\frac{\partial \phi^{OC}}{\partial P^k}$ is this average partial effect on ϕ^{OC}.

$$(11)\quad \left.\frac{\partial \mathbb{P}}{\partial P^k}\right|_{SR} = \underbrace{\mathbb{P}\frac{\partial \phi^{CC}}{\partial P^k}}_{\left.\frac{\partial \mathbb{P}^{CC}}{\partial P^k}\right|_{SR}} + \underbrace{(1-\mathbb{P})\frac{\partial \phi^{OC}}{\partial P^k}}_{\left.\frac{\partial \mathbb{P}^{OC}}{\partial P^k}\right|_{SR}}$$

The long-run marginal effect allows influence from changes in the steady-state probability as well as the transitory effects of the transition matrix. Equation 12 shows the long-run marginal effect of a price change and can be derived by inserting the derivative of equation 9 multiplied by $\left(\phi^{CC}-\phi^{OC}\right)$, as a second term. Like the short-run marginal effects, the long-run state marginal effects will be a sum of the long-run marginal effects of prices on the sequence probabilities $\mathbb{P}^{OC}$ and $\mathbb{P}^{CC}$.

$$(12)\quad \left.\frac{\partial \mathbb{P}}{\partial P^k}\right|_{LR} = \underbrace{\left.\frac{\partial \mathbb{P}}{\partial P^k}\right|_{SR} + \left[\phi^{CC}-\phi^{OC}\right]\frac{\left[1-\phi^{CC}\right]\frac{\partial \phi^{OC}}{\partial P_k} + \phi^{OC}\frac{\partial \phi^{CC}}{\partial P_k}}{\left[1-\phi^{CC}+\phi^{OC}\right]^2}}_{\left.\frac{\partial \mathbb{P}^{CC}}{\partial P^k}\right|_{LR} + \left.\frac{\partial \mathbb{P}^{OC}}{\partial P^k}\right|_{LR}}$$

For interpretation, it is important to emphasize what is meant by long and short-run estimates. The long-run estimates refer to the marginal effects of each of the regressors on a dynamic process that has reached a steady-state equilibrium. At the steady state equilibrium, producers have a consistent probability of transitioning to corn year after year. The short-run effect refers to the immediate impact that a price has on the probability of transitioning to corn. In this case, the probability of planting corn has not reached an equilibrium and is conditional on the lagged crop choice as opposed to the steady state probability. We included a section further explaining the difference between long- and short-run effects in the supplementary online appendix.

We can estimate the effect that prices have on rotations using the sequential marginal effects. For continuous crop rotations, the sequential marginal effects equal the rotational effects. This is because individuals in the continuous corn (other) rotation perform a corn-corn (other-other) planting sequence each year. This is not true for those in the other-corn rotation. Half of the time, individuals in these rotations perform an other-corn sequence and the other half they perform a corn-other sequence. Therefore the probability of an other-corn rotation is the average of these sequential probabilities. Equations 13, 14, and 15 show the relations between each rotational probability and the sequential probabilities.

$$\mathbb{P}^{\{CC\}ROT} = \mathbb{P}^{CC} \tag{13}$$

$$\mathbb{P}^{\{OO\}ROT} = \mathbb{P}^{OO} \tag{14}$$

$$\mathbb{P}^{\{OC\}ROT} = \frac{1}{2}\left[\mathbb{P}^{OC} + \mathbb{P}^{CO}\right] \tag{15}$$

Aggregate Elasticities

In this section we discuss how we translate our field-level marginal effects to elasticity terms and how we aggregate these terms to the national level. We construct our acreage-weighted elasticities using an acreage weighted sum of our field level marginal effects divided by the expected corn acreage in the country and multiply by the respective crop's national average price. We represent the expected national corn acreage, the numerator term, by summing each field's acreage multiplied by its expected corn planting probability. Similarly, we estimate the quantity-weighted elasticities by weighting each term by their respective observed NASS county yields and field size.

Equation 16 shows the aggregate acreage-weighted elasticity.

$$(16)\quad \varepsilon^{Acre} = \left(\sum_{i=1}^{N} \frac{\partial \mathbb{P}_i}{\partial P^k} \times acre_i \right) \frac{\bar{P}^k}{\sum_{i=1}^{N} \mathbb{P}_i \times acre_i}$$

Here, P^k is the price of crop k, $\bar{P}^k$ is the national average crop k price, and $acre_i$ is the acreage of field i. Equation 17 shows the quantity-weighted elasticity where $yield_i$ is the county-level yield for the respective field i provided by the National Agricultural Statistics Service (NASS).

$$(17)\quad \varepsilon^{Qty} = \left(\sum_{i=1}^{N} \frac{\partial \mathbb{P}_i}{\partial P^k} \times acre_i \times yield_i \right) \frac{\bar{P}^k}{\sum_{i=1}^{N} \mathbb{P}_i \times acre_i \times yield_i}$$

Our marginal effect terms for each field are derived from MLRA-soil groups. That is, every field within these groups has a common marginal effect. It may therefore seem odd that we aggregate these results at the field and not at the group-level. However, in both of these elasticity terms, we use the predicted probability terms at the field level $\mathbb{P}_i$. This is an important distinction since these predicted probabilities are a function of field-level characteristics. To account for this, we aggregate from the field-level. As a result of this weighting, the average elasticities depart from their marginal effect counterparts.

Standard Error Calculation

We account for the correlation in the data by clustering by year, which assumes that the unobserved determinants of crop choice are independent between years but allows for spatial correlation within each year.[11] Estimating standard errors on data with a small number of clusters (generally less than 30) is challenging as many methods such as sandwich estimators and paired clustered bootstraps rely on large samples to construct unbiased estimates of standard errors (Cameron and Miller, 2015; Cameron, Gelbach, and Miller, 2008). Cameron, Gelbach, and Miller (2008) examine the issue of estimating standard errors in datasets with few clusters and find that the wild bootstrap performs well relative to other resampling methods. Unlike the standard residual bootstrap, the wild bootstrap preserves the distribution of the original error terms. This is an important property if the distributions of error terms vary across clusters. The wild bootstrap creates pseudo-error terms by multiplying the errors from the original regression by some random variable with a mean of zero and a unit standard deviation. This ensures that the pseudo-error terms have the same distribution as the original error terms. In the wild bootstrap, the pseudo-error terms are added to the corresponding dependent variable estimates to produce pseudo-independent variables which are used to obtain a new set of coefficients.

Since models like the logit are estimated via maximum likelihoood and do not produce observed errors, the standard wild bootstrap is not practical in our application. Instead, we use the wild score bootstrap from Kline and Santos (2012) and recommended by Cameron and Miller (2015). To compute cluster robust statistics using the wild score bootstrap, we apply a common perturbation value W_i to every observation within each cluster. To simplify the wild bootstrap, the random perturbation W_i usually takes on discrete values. While the Rademacher distribution and the distribution described by Mammen (1993) are popular choices for W_i, they only take on two discrete values. This would only provide ($2^9 = 512$) possible randomized samples in the most data restricted MLRA-groups observed between 2008 and 2016. We instead use an alternative weighting scheme proposed by Webb (2013) which incorporates a discrete, uniformly distributed, random variable W_i taking on the values $\left\{-\sqrt{\frac{3}{2}}, -\sqrt{\frac{2}{2}}, -\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}}, \sqrt{\frac{2}{2}}, \sqrt{\frac{3}{2}}\right\}$ with $\frac{1}{6}$ probability. This method provides a sampling procedure while increasing the number of potential random samples when there are only a few clusters (Webb, 2013). As a result, Webb's distribution expands the potential random samples of the nine-year clustered bootstrap to ($6^9 = 10,077,696$) and is consequentially more appropriate for bootstrapping.

The primary difference between the wild score bootstrap and the standard wild bootstrap is that it is applied to score function components instead of error terms. Here y_i is the dependent binary variable, x_i is a $(1 \times k)$ vector of independent variables, and β is a $(1 \times k)$ vector of coefficients where (i) indexes observations. The wild score bootstrap creates coefficient replications by perturbing the original score function components by W_i and performing an additional Newton-Raphson iteration using the perturbed score function, shown in equation 18.

$$(18) \quad \hat{\beta}^{wild} = \hat{\beta} - \mathbb{H}^{-1}\frac{1}{n}\sum_i \left[y_i - \left(1 + \exp\left\{-x_i'\hat{\beta}\right\}\right)^{-1} \right] x_i W_i$$

Here (n) is the total number of observations, $\left(\hat{\beta}\right)$ is the original coefficient vector and $(-\mathbb{H})$ is the Fischer Information matrix. This bootstrap procedure has the advantage that once the original model is estimated, it does not require re-estimation of the model. The wild score bootstrap also avoids inverting the Fischer Information matrix which dramatically reduces the computational burden of bootstrapping (Kline and Santos, 2012). This is an attractive feature of the score bootstrap as it makes carrying out the 1,000 iteration bootstrap over our 32 million observations tractable.

Results

In this section, we show the findings of our models. We discuss how corn planting transitions interact with our controls, how corn planting and rotations respond to price changes, how these responses vary across the country, and compare the aggregate results of our model that allows for heterogeneous response against one that imposes a uniform planting response.

Markov Transition Equation Coefficient Summary

We start by presenting a summary of the coefficient distributions across different areas of the country. While coefficient values in the logit models do not correspond to marginal effects, their signs and standard errors are consistent with the signs of the corresponding marginal effects. In later sections we consider the marginal effects and elasticities with respect to prices. We focus on the coefficients here due to the computational demands of constructing the distribution of marginal effects for every variable in our model.[12] Figure 2 illustrates the coefficients on the independent variables in the transition probabilities using sets of whisker and dot plots.[13] The dots show the coefficient percentile estimates at the 10%, 25%, 50%, 75%, and 90% levels across the 115 MLRA soil groups. For example, the red dot for the 90th percentile for corn price shows the coefficient estimate on corn price in the ϕ^{OC} equation that is larger than 90% of the other coefficients across the 115 groups. These dots approximate the cumulative distributions of the coefficient estimates across the country for each model. The line around the dot indicates the 95% confidence interval of the estimate constructed using 1,000 bootstrapped estimates. The lack of overlap across the confidence intervals suggest that the coefficient heterogeneity is statistically significant between regions.

The price coefficients largely conform to economic theory. Most of the corn price coefficients were positive in both the ϕ^{OC} and ϕ^{CC} models. Price coefficients at the 10th percentile in both models were near zero. The other price coefficients were negative for 75% of the groups.

There is more heterogeneity on the sign of the impact of soil characteristics across groups. The median of the field slope coefficient in both models was approximately zero. This means that around half of the coefficients are positive and the other half are negative. However, the impact of soil productivity on transitioning to corn when corn was previously planted varies across regions. In some regions, more productive soils increase the probability of planting corn, but in many regions higher productivity soil decreases the probability of planting corn when corn was previously planted. This suggests that relative returns to growing corn on better quality land differ across the country. There are many reasons these returns could differ including regional differences in viable alternative crops to corn, local markets, and climatological features.

Pre-plant weather also has varied effects across the country. Dry pre-plant conditions produced diverse effects. The effects are symmetric with a median near zero. In some areas of the country, low pre-plant precipitation reduces the probability of transitioning to corn while increasing the probability in other areas. Wet conditions largely reduce the probability of transitioning to corn across the country. Wet pre-plant conditions provide a host of problems including reduced yields, soil compaction, and planting delays also lead to yield penalties (MacKellar and Anderson, 2016; Farnham, 2001). The major climatological variation across the country likely contributes to the lack of consistency in the precipitation effects. Exceptionally wet years in dry areas may be optimal for corn production, while exceptionally wet years in wet areas could leave the soil unsuitable for planting.

Irrigated fields are more likely to transition to corn relative to non-irrigated fields. Corn has demanding evapotranspiration requirements and corn yields are especially sensitive to drought (Stone and Schlegel, 2006; Leng and Hall, 2019). It is therefore not surprising that transitions to corn are more probable on irrigated fields. The time trends differ across models, the ϕ^{CC} models has negative trends and the ϕ^{OC} models has positive trends. This suggests that, holding all else constant, corn-corn transitions are becoming less popular and other-corn transitions are becoming more popular over time.

Planted Acreage Elasticities

We now move to the primary focus of the article, the relationship between intensive corn planting and prices. Figure 3 illustrates the short- and long-run corn planting elasticities at the 10th, 25th, 50th, 75th and 90th percentiles across the country.[14] As in the previous figure, the dots represent the means of elasticity percentiles and the whiskers represent their 95% confidence intervals. The corn acreage elasticities conform to economic theory. The

10th percentile statistics are all close to zero. This indicates that rotations in 10% of the areas in this study are not responsive to price changes. The own-price corn planting elasticity is generally positive as approximately 75% of the groups have inelastic own-price planting responses. Over 10% of the areas have elastic own-price planting responses.

The cross-price corn planting elasticities mirror the own-price elasticities. As expected, these elasticities are negative which indicates that an increase in the other price index reduces the planted corn acreage. While around 25% of the areas have elasticities at or near zero, the majority of areas have inelastic planting responses to other prices. Around 10% of the areas have elastic planting responses to other crop prices.

While informative, these percentiles are not weighted by historic corn production. Figure 4 shows the cumulative share of corn production from 2009 to 2016 by the own-price elasticity estimates. This shows that a little over 20% of the corn in the country was planted in areas with elastic short-run own-price planting response and around 20% of total corn production occurred in areas with elastic long-run planting response.

Planting Response Across Space

Our analysis allows for price response comparisons across the entire country by estimating separate effects for each MLRA-soil texture group. Figures 5 and 6 show the field-level own- and cross-price marginal effects on planted corn acreage. These results show the spatial pattern in planting response across the country. Planting inside the traditional Corn Belt was moderately sensitive to corn prices. In states like Iowa, Illinois, Indiana, and Nebraska, a $1 increase in the price of corn increases planted corn acreage by approximately 10 percentage points. Corn plantings in the eastern Dakotas, western Minnesota, southern Wisconsin, central Michigan, and the Mississippi River Delta are more sensitive to price. In some of these areas, a $1 increase corn price increases corn plantings by 30 percentage points. However, there are areas outside of the Corn Belt that are not very sensitive with the notable examples of Kansas, and the East Coast. In some areas, local growing conditions may constrain planting response. For example, western Kansas is prone to droughts, relies heavily on irrigation, and has less productive soil in the southern half of the state. Marketing outlets are another consideration. Corn planting in the southeastern states also is relatively unresponsive to price changes. This an important area for the broiler industry. If most of the corn local supply serves as the local animal feed, planting decisions may be less responsive to general price movements. While there are many potential causes for the diversity in the price responsiveness in planting, the primary focuses of this article is to characterize and quantify the level of price response heterogeneity and not necessarily to diagnose its underlying causes. We leave this to future research.

The moderate price response in the traditional corn states such as Nebraska, Iowa, Indiana, and Illinois arises because the most popular crop rotations in these states already include corn. Crop production on the fringes of the Corn Belt is more diverse and corn is a smaller proportion of total production. As such, areas outside of the Corn Belt may suffer lower yield penalties if they increase corn production on the intensive margin. Persistent basis patterns may be another explanation for high sensitivities in the North Central US. Corn prices in the eastern Dakotas and western Minnesota tend to be lower relative to the rest of the country which may make growing corn unprofitable without favorable movements in price. If true, planting in the region may respond more aggressively than in areas where corn prices are already high. Researchers have noted high price sensitivity among crop producers in the Eastern Dakotas (Wang et al., 2017). Our approach of separately modeling different areas of the country would take this into account as basis patterns are generally consistent over space.

These maps also show that there is a degree of within-MLRA differences in price sensitivity. The MLRA on the border of the eastern Dakota states and Minnesota provides a clear example of within-MLRA heterogeneity by soil texture type. The Red River of the North provides the boundary for these states. Soils near the Red River contain more clay and are more susceptible to flooding in early springs. This could be a reason that corn plantings are less sensitive to price, than those farther away from the river.

Rotational Response

While the influence that prices have on planting corn is important when estimating supply elasticities, their effect on crop rotations can have important implications on their own. Continuously planting nitrogen intensive crops tends to increase costs, yield variability, and negative environmental impacts. Figure 7 shows the 10th, 25th, 50th, 75th and 90th percentiles for the corn-corn, other-other, and other-corn rotation probability elasticities with respect to changes in the corn price and the other price index across the 115 groups. Again, the signs of the elasticity terms are consistent with economic theory. Higher corn prices increase the likelihood of continuous corn rotations, decrease the probability of continuous other rotations, and have mixed influence on other-corn rotations. Increases in the other price index produced opposite effects. Measured as elasticities, the rotation selection responses are far more elastic than the corn probability response. The continuous rotation choice in approximately half of the areas is elastic to price. Continuous cropping has implications for future yields and raises environmental concerns. These results suggest that policies that support corn prices could significantly impact the environment via rotational response.

Unlike the planted corn elasticities, the short-run rotational elasticities are smaller in magnitude than their long-run counterparts. This is because rotations are a multi-year concept. It makes sense that a temporary fluctuation in prices would not have a large influence over multi-seasonal decisions of farmers relative to a more persistent price change. This supports the results and discussion of Hendricks et al. (2014) who found that corn planting elasticities in the short-run where larger than their long-run counterparts while the opposite was true for rotation elasticities. There were mixed price effects on other-corn rotation across the country. The median other-corn rotation elasticity was close to zero. The small effect of prices on other-corn crop rotations is not surprising since, on average, these rotations plant corn half the time and other crops half the time.

Comparing Heterogeneous and Pooled Models

Preventing bias is one of the primary motivations for estimating the model using heterogeneous coefficients. To illustrate the benefits of incorporating heterogeneity, we estimate an alternative model where we estimate each Markov transition equation once over the pooled national sample. The specification of the pooled model is nearly identical to the heterogeneous models. We use the same price variables in the pooled model, but we also include MLRA fixed-effects in the pooled model since we constructed the other price index using MLRA-level production baskets. The MLRA fixed-effects correct for the fact that the other price index is endogenous between MLRAs since it depends on the crops produced in each MLRA. Without these MLRA fixed-effects, the pooled model results are inconsistent with economic theory (e.g., increasing corn prices decreases corn acreage). By estimating a single model, we compare how the national-level rotation-related supply elasticity estimates differ when we allow for rotational response heterogeneity.

Table 3 provides a summary of national-level elasticity estimates from heterogeneous and pooled (i.e., uniform rotational response) models. The first result is that the acreage-weighted supply response is larger than the quantity-weighted response. This finding is consistent with a Ricardian model where production on marginal land (i.e., less productive land) becomes more profitable as prices increase (Ricardo, 1821). The Corn Belt-only acreage and quantity weighted elasticities are much closer to one another. This is likely due to more behavioral and yield homogeneity within the region. Fields that are closer to one another likely have similar yields and production practices.

Table 3 also illustrates the differences between the elasticities from the national heterogeneous response and the national pooled models. National elasticities that allow for heterogeneous responses are larger than the pooled estimates. Both the short- and long-run differences are substantial. The pooled model estimates are between 11 to 19 percentage points smaller than elasticities from the heterogeneous model. The results also indicate that the pooled model does not accurately estimate the dynamics of the response.

Table 3 also shows the relative difference between the long- and short-run elasticities $\left(\frac{\varepsilon_{LR}-\varepsilon_{SR}}{\varepsilon_{LR}}\right)$ between these models. We refer to this term as the lagged dependent term because it is analogous to the coefficient on the lagged dependent variable in a single equation autoregressive model. When this term is close to zero, it means that the prior crop choice does not influence the current one. When it is near -1, producers tend to alternate crops every year. The lagged dependent terms are negative since farmers generally follow non-continuous rotations. The lagged dependent terms in the pooled model are generally smaller than those in the heterogeneous model. This means that the pooled model underestimates the tendency to rotate away from crops planted in the prior season.

Table 4 shows the statistical properties of the differences between the national heterogeneous response model and the national-level response models. Because, we use consistent wild score bootstrap weights between the heterogeneous response and pooled models across our bootstraps for each model, we can construct a distribution of the differences between these average elasticities directly by differencing these values from each bootstrap iteration. As expected, the average differences closely correspond to the results in table 3. The p-values, particularly on the elasticity differences were quite small with all being statistically different from zero at or near the 1% level. This shows that heterogeneous model estimates provide a strong statistical improvement in elasticity estimates relative to models that do not account for possibility of different price effects across the country.

Table 4 also shows the differences between the lagged dependent variables for the own and cross price elasticities. The lagged dependent variables from the own-price elasticities were significant at or near the 10% level. This provides evidence that failing to allow for heterogeneous price effects understates the lagged dependent variable by between three to six percentage points, though the significance of this bias is weaker than the bias of the elasticities. The sign of this bias aligns with the theoretical predictions of Pesaran and Smith

(1995). The differences between the models in the lagged dependent variable in the cross price elasticities is much less pronounced and not consistent across acreage and quantity weightings. This may be due to our broad definition of the "other" crop category as a given rotation pattern could involve many different "other" crops regardless of the modeling assumptions.

The last important result from table 3 is that the national-level elasticities are much more elastic than in the Corn Belt. One reason for a more elastic response from areas outside the Corn is that less of the increase in corn acres comes from an expansion in continuous cropping. Table 5 shows transition and sequential probabilities for the different pairs of planting decisions. Recall the sequential probabilities for corn-corn and other-other also represent their rotation probabilities. The corn-other rotation probability is the sum of the corn-other and other-corn sequential probabilities. In the case of the Corn Belt, the corn-other rotation is extremely popular. Over 75% of the Corn Belt observations are in corn-other rotations while only 58% are in corn-other rotations outside the Corn Belt. This means that the majority of the potential response within the Corn Belt will come from fields in corn-other rotations. Continuous other rotations are also more popular outside of the Corn Belt. Fields inside the Corn Belt have only an 8% probability of performing a continuous other rotation while this probability is nearly 30% those outside the Corn Belt. This means that greater proportions of the area outside of the Corn Belt can adjust their corn rotations by transitioning from a continuous other rotation to corn-other rotations. Since more of the possible response outside the Corn Belt can comes from adopting corn-other rotation, as opposed to corn-corn rotations, corn rotations in areas outside of the Corn Belt may be more responsive to prices because they can better avoid yield drag and increased input expenditures associated with adopting continuous corn rotations.

Differences in the conditional transitional probabilities between the regions may be another reason for the differences in corn rotation responsiveness. The conceptual model in Wang et al. (2017) suggests that the rotational response will be larger in areas where the conditional probability of a binary planting decision is close to 50%. The logic is that if a producer has the choice between corn and some other crop, and the probability of planting corn is close to 50%, conditional on the previous crop choice, then the probability of some other choice is also close to 50%. Probabilities around 50% are signs of indecisiveness and even small price changes are more likely to affect crop transitions. The leading term in logit marginal effects of the transition probabilities $\left(\phi^{CC}\left(1-\phi^{CC}\right)\beta_1^k\right)$ incorporate this idea because the marginal effect is largest when the transition probability is 50%.

Even if the coefficients on price were the same inside and outside the Corn Belt, we would expect differences in the effects of prices purely due to differences in the transition probabilities. Specifically, under homogeneous coefficients, the own-price marginal effect on other-corn sequences is 2.16 times larger outside the Corn Belt. While the own-price marginal effect on corn-corn sequences is 0.74 times as large outside the Corn Belt.[15] Therefore, the overall own-price marginal effect on the probability of planting corn is likely to be much larger outside the Corn Belt since the marginal effect on other-corn is so much larger even with the same coefficients. This highlights the importance of considering areas where corn is viable and not just where corn is popular. This does not explain all of the differences in price sensitivity because coefficients also vary significantly across the country as we showed in figure 2.

Conclusions

Our results show that the rotational component of total national corn supply elasticities are higher than the existing literature suggests. Many of the studies on the welfare impacts of trade and other policy interventions rely on supply elasticities for accurate estimates. We separately model corn supply response across 115 subregions of the country. We find a high degree of supply response heterogeneity. While estimating environmental impacts is not our aim, understanding supply response heterogeneity is important when predicting the environmental impacts from changing prices. We find that the Prairie Pothole Region has some of the most responsive corn supply in the country. Our results are especially useful for environmental studies since we estimate rotational elasticities, which can have distinct environmental implications. These results suggest continuous corn rotations are largely elastic across the country. Our results also show that corn supply and acreage are more sensitive in the short-run than in the long-run consistently across the country. This suggests that the benefits of rotations play an important role in supply response to prices and generalizes the findings from Hendricks, Smith, and Sumner (2014) to the national-level.

The dataset in this study is unique in the literature and includes crop choices at the field-level that account for over 83% of the land devoted to corn production in the US between 2009 and 2016 both in and out of the traditional Corn Belt states. Our results show that modeling supply response without accounting for heterogeneity can bias the national estimates. Supply in the Corn Belt tends to be less sensitive than the national estimates. Corn rotations are especially elastic in Northern Plains states and the Mississippi Delta and insensitive in western Kansas and the Gulf states. This cautions against extrapolating such estimates extra-regionally.

In this analysis, we group fields by MLRA and soil texture group to incorporate heterogeneity but maintain large enough samples for model stability within each area. The results from Pesaran and Smith (1995) show that pooling bias shrinks as coefficient heterogeneity declines. The statistically significant heterogeneity in the planting response between MLRA-groups provides evidence that our aggregate estimates are an improvement over pooled estimates. However, this does not preclude further estimation improvements from alternative groupings. The optimal grouping strategy is left as a topic for future research.

References

Barr, K.J., B.A. Babcock, M.A. Carriquiry, A.M. Nassar, and L. Harfuch. 2011. "Agricultural Land Elasticities in the United States and Brazil." *Applied Economic Perspectives and Policy* 33:449–462.

Benham, E.C., R.J. Ahrens, and W.D. Nettleton. 2009. "Clarification of Soil Texture Class Boundaries." MO5 Soil Technical Note No. 16, National Soil Survey Center, USDA-NRCS, March.

Berry, S., and W. Schlenker. 2011. "Technical Report for the ICCT: Empirical Evidence on Crop Yield Elasticities." *International Council on Clean Transportation. http://www. theicct. org/sites/default/files/publications/berry_schlenker_cropyieldelasticiti es_sep2011. pdf*, pp. 1–19.

Brown, J.F., and M.S. Pervez. 2014. "Merging Remote Sensing Data and National Agricultural Statistics to Model Change in Irrigated Agriculture." *Agricultural Systems* 127:28–40.

Cameron, A.C., J.B. Gelbach, and D.L. Miller. 2008. "Bootstrap-Based Improvements for Inference With Clustered Errors." *The Review of Economics and Statistics* 90:414–427.

Cameron, A.C., and D.L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources* 50:317–372.

Claassen, R., C. Langpap, and J. Wu. 2017. "Impacts of Federal Crop Insurance on Land Use and Environmental Quality." *American Journal of Agricultural Economics* 99:592–613.

Davis, G.C., and M.C. Espinoza. 1998. "A Unified Approach to Sensitivity Analysis in Equilibrium Displacement Models." *American Journal of Agricultural Economics* 80:868–879.

Dobos, R., H. Sinclair, and K. Hipple. 2008. "User Guide National Commodity Crop Productivity Index (NCCPI) Version 1.0." *US Department of Agriculture, Natural Resources Conservation Service*, pp. 1–69.

Donaldson, D., and A. Storeygard. 2016. "The View from Above: Applications of Satellite Data in Economics." *Journal of Economic Perspectives* 30(4):171–98.

Farnham, D.E. 2001. "Corn Planting Guide." Extension Article No. PM 1885, Iowa State University.

Gardner, B.L. 1976. "Futures Prices in Supply Analysis." *American Journal of Agricultural Economics* 58:81–84.

Haile, M.G., M. Kalkuhl, and J. Braun. 2014. "Inter-and Intra-Seasonal Crop Acreage Response to International Food Prices and Implications of Volatility." *Agricultural Economics* 45:693–710.

Haile, M.G., M. Kalkuhl, and J. von Braun. 2016. "Worldwide Acreage and Yield Response to International Price Change and Volatility: A Dynamic Panel Data Analysis for Wheat, Rice, Corn, and Soybeans." *American Journal of Agricultural Economics* 98:172–190.

Hendricks, N.P., J.P. Janzen, and A. Smith. 2014. "Futures Prices in Supply Analysis: Are Instrumental Variables Necessary?" *American Journal of Agricultural Economics* 97:22–39.

Hendricks, N.P., S. Sinnathamby, K. Douglas-Mankin, A. Smith, D.A. Sumner, and D.H. Earnhart. 2014. "The Environmental Effects of Crop Price Increases: Nitrogen Losses in the US Corn Belt." *Journal of Environmental Economics and Management* 68(3):507–526.

Hendricks, N.P., A. Smith, and D.A. Sumner. 2014. "Crop Supply Dynamics and the Illusion of Partial Adjustment." *American Journal of Agricultural Economics* 96:1469–1491.

Hennessy, D.A. 2006. "On Monoculture and the Structure of Crop Rotations." *American Journal of Agricultural Economics* 88:900–914.

Kim, H., and G. Moschini. 2018. "The Dynamics of Supply: US Corn and Soybeans in the Biofuel Era." *Land Economics* 84:593–613.

Kline, P., and A. Santos. 2012. "A Score Based Approach to Wild Bootstrap Inference." *Journal of Econometric Methods* 1:23–41.

Koutchadé, O.P., A. Carpentier, and F. Femenia. 2018. "Modeling Heterogeneous Farm Responses to European Union Biofuel Support with a Random Parameter Multicrop Model." *American Journal of Agricultural Economics* 100:434–455.

Lacroix, A., and A. Thomas. 2011. "Estimating the Environmental Impact of Land and Production Decisions with Multivariate Selection Rules and Panel Data." *American Journal of Agricultural Economics* 93:784–802.

Langpap, C., and J. Wu. 2011. "Potential Environmental Impacts of Increased Reliance on Corn-Based Bioenergy." *Environmental and Resource Economics* 49:147–171.

Lee, D.R., and P.G. Helmberger. 1985. "Estimating Supply Response in the Presence of Farm Programs." *American Journal of Agricultural Economics* 67:193–203.

Leng, G., and J. Hall. 2019. "Crop Yield Sensitivity of Global Major Agricultural Countries to Droughts and the Projected Changes in the Future." *Science of the Total Environment* 654:811–821.

MacKellar, B., and E. Anderson. 2016. "Planting Corn in Wet Conditions – Is it Worth it?" Extension Article, Michigan State University Extension.

Mammen, E. 1993. "Bootstrap and Wild Bootstrap for High Dimensional Linear Models." *The Annals of Statistics* 21:255–285.

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

Moschini, G., H. Lapan, and H. Kim. 2017. "The Renewable Fuel Standard in competitive equilibrium: Market and welfare effects." *American Journal of Agricultural Economics* 99:1117–1142.

Motamed, M., L. McPhail, and R. Williams. 2016. "Corn Area Response to Local Ethanol Markets in the United States: A Grid Cell Level Analysis." *American Journal of Agricultural Economics* 98:726–743.

NASS, USDA. 2010. "Field Crops: Usual Planting and Harvesting Dates." *USDA National Agricultural Statistics Service, Agriculural Handbook* 628.

Nerlove, M. 1956. "Estimates of the Elasticities of Supply of Selected Agricultural Commodities." *Journal of Farm Economics* 38:496–509.

Orazem, P.F., and J.A. Miranowski. 1994. "A Dynamic Model of Acreage Allocation With General and Crop-Specific Soil Capital." *American Journal of Agricultural Economics* 76:385–395.

Pervez, M.S., and J.F. Brown. 2010. "Mapping Irrigated Lands at 250-m Scale by Merging MODIS Data and National Agricultural Statistics." *Remote Sensing* 2:2388–2412.

Pesaran, M.H., and R. Smith. 1995. "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels." *Journal of econometrics* 68:79–113.

Ricardo, D. 1821. *Principles of Political Economy and Taxation*, 3rd ed. London: John Murray.

Roberts, M.J., and W. Schlenker. 2013. "Identifying Supply and Demand Elasticities of Agricultural Commodities: Implications for the US Ethanol Mandate." *The American Economic Review* 103:2265–2295.

Stone, L.R., and A.J. Schlegel. 2006. "Crop Water Use in Limited-Irrigation Environments." In *Proceedings of the 2006 Central Plains Irrigation Conference, Colby, KS*. Citeseer, pp. 21–23.

Tegene, A., W.E. Huffman, and J.A. Miranowski. 1988. "Dynamic Corn Supply Functions: A Model With Explicit Optimization." *American Journal of Agricultural Economics* 70:103–111.

Wang, S.L., P. Heisey, D. Schimmelpfennig, and V.E. Ball. 2015. "Agricultural Productivity Growth in the United States: Measurement, Trends, and Drivers." *Economic Research Service*, July, pp. 1–68.

Wang, T., M. Luri, L. Janssen, D.A. Hennessy, H. Feng, M.C. Wimberly, and G. Arora. 2017. "Determinants of Motives for Land Use Decisions at the Margins of the Corn Belt." *Ecological Economics* 134:227–237.

Webb, M.D. 2013. "Reworking Wild Bootstrap Based Inference for Clustered Errors." Working paper, Queen's Economics Department Working Paper.

Woodard, J. 2016. "Ag-Analytics.org Data Warehouse." Available online at: ag-analytics.org.

Wu, J., R.M. Adams, C.L. Kling, and K. Tanaka. 2004. "From Microlevel Decisions to Landscape Changes: An Assessment of Agricultural Conservation Policies." *American Journal of Agricultural Economics* 86:26–41.

Yan, L., and D. Roy. 2016. "Conterminous United States Crop Field Size Quantification from Multi-Temporal Landsat Data." *Remote Sensing of Environment* 172:67–86.

Notes

[1]See tables A1 and A2 in the online appendix for a summary of the 2016 Cropland Data Layer classification accuracy rates across each of the states in the analysis by crop.

[2]Note that if X, Y, and Z are random variables, then $\mathbb{E}[XYZ] = Cov(X, YZ) + \mathbb{E}[X]\mathbb{E}[YZ]$.

[3]For a complete map of the available years of CDL data by MLRA, see table A4 in the online appendix.

[4]For a description of how we define soil class "closeness" see the "Combining Soil Texture Groups" section in the online appendix.

[5]For a complete list of crop definitions see table A5 in the online appendix.

[6]Bloomberg price series were accessed on November 7th 2017. To complete the data for some sites, we placed an ad hoc price series order directly from Cash Grain Bids. These ordered data were accessed on February 6th 2018.

[7]The computation of this crop basket assumes that every crop grown in an area from 2008 to 2016 was a relevant crop over the entire period. While technological changes could introduce some of these crops over time, incorporating technological trends into the production basket weakens the effectiveness of the price index since temporal variation in the index will not entirely be from price variation.

[8]Corn planting for many of the largest corn-producing states is most active in these months (NASS, USDA, 2010).

[9]Note that sequences are not the same as rotations. While rotations are repeated, sequences need not be. For instance, a crop sequence $\{oth, crn\}$ occurs when a farmer plants an "other" crop and then plants corn in the following period. An $\{oth, crn\}$ rotation is when a farmer plants some other crop in the previous period and corn in the following period *and then* immediately repeats this sequence. If a farmer adopts the $\{oth, crn\}$ rotation, half of the time, she will plant an $\{oth, crn\}$ sequence and the other half he will plant a $\{crn, oth\}$ sequence. Conversely, continuous rotations (e.g. the $\{crn, crn\}$ and $\{oth, oth\}$ rotations) are entirely composed of their respective continuous sequences and are therefore identical to their respective sequences.

[10]The steady-state probability of planting other crops is the complement of the steady-state probability of planting corn.

[11]See figures A2 and A3 in the online appendix for a graphical representation of this.

[12]Considering the distributions of the marginal effects would require 6 additional 1,000 bootstraps across 230 distinct models with a dataset of more than 30 million observations would require substantial computing time, even on the Beowulf cluster that we used.

[13]See tables A6 A7, and in the online appendix for numerical summaries of the coefficient values.

[14]See the table in the online appendix for a numerical summary of the elasticity values.

[15]This follows from the components in equation 11 , the functional form of logit marginal effects, and the values from table 5. From this, if the coefficients are identical between the two areas, the sequential marginal effect ratios for other-corn sequence is $\frac{\frac{\partial \mathbb{P}^{OC}}{\partial P_k}|_{outside}}{\frac{\partial \mathbb{P}^{OC}}{\partial P_k}|_{inside}} = \frac{0.495(1-0.495)}{0.817(1-0.817)}\frac{(0.289+0.294)}{(0.369+0.082)} = 2.16$. The ratios of sequential marginal effects for corn-corn sequences is $\frac{\frac{\partial \mathbb{P}^{CC}}{\partial P_k}|_{outside}}{\frac{\partial \mathbb{P}^{CC}}{\partial P_k}|_{inside}} = \frac{0.307(1-0.307)}{0.327(1-0.327)}\frac{(0.128+0.289)}{(0.179+0.369)} = 0.74$.

Figures

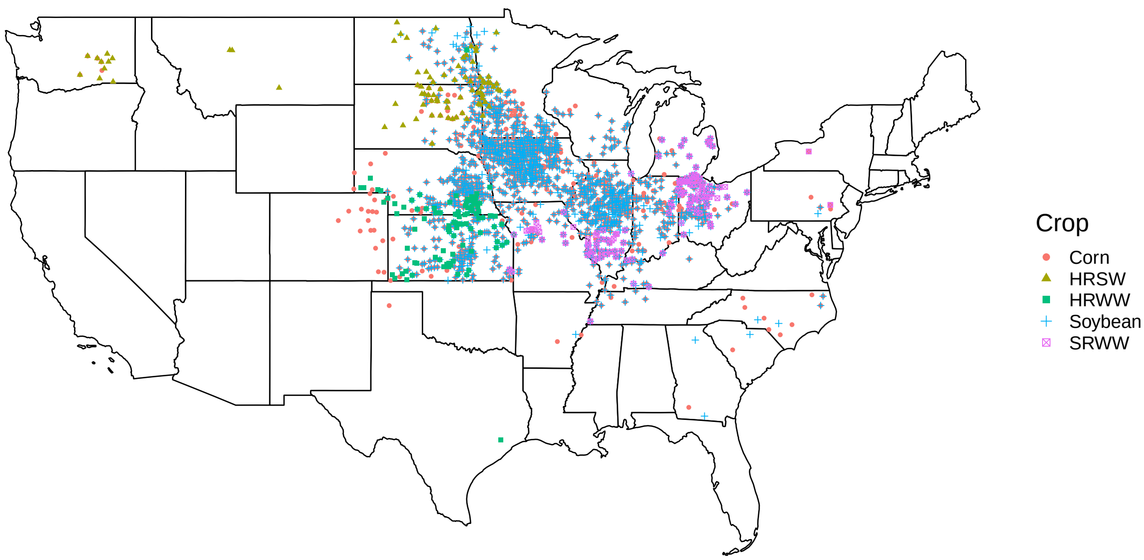
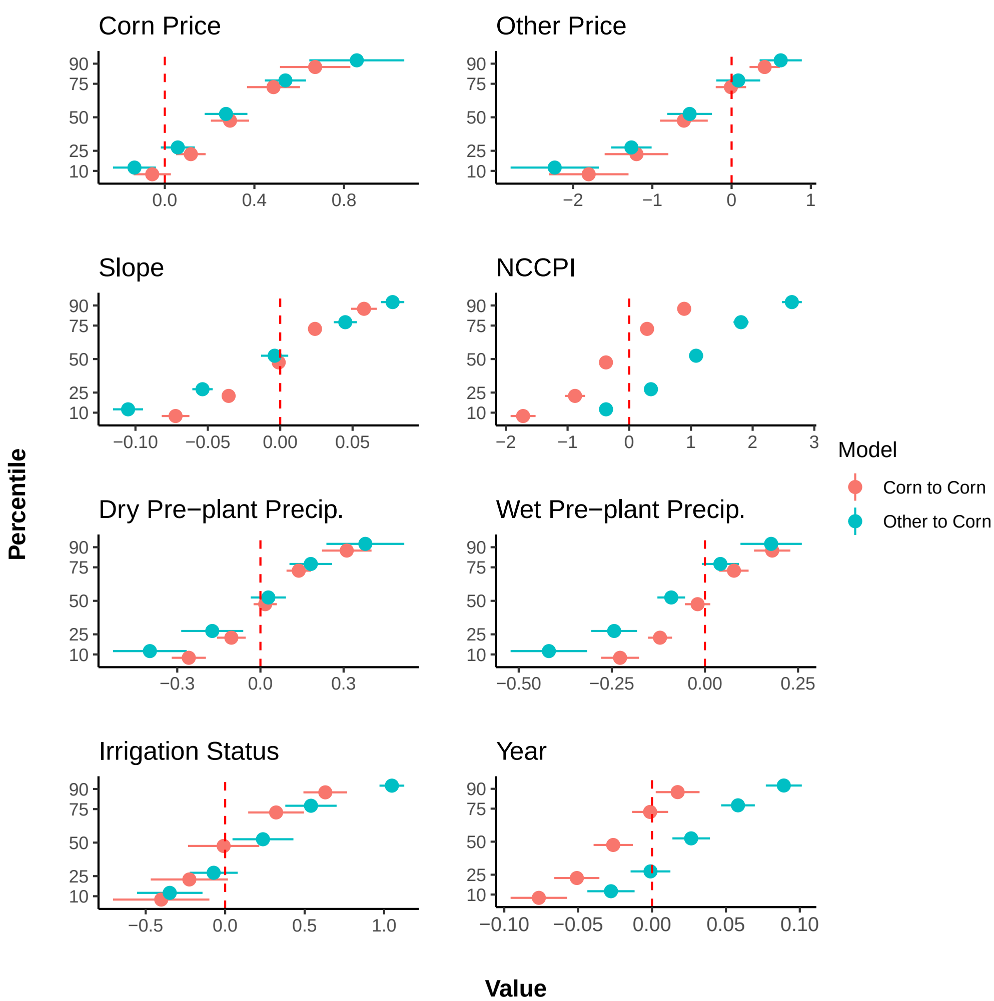
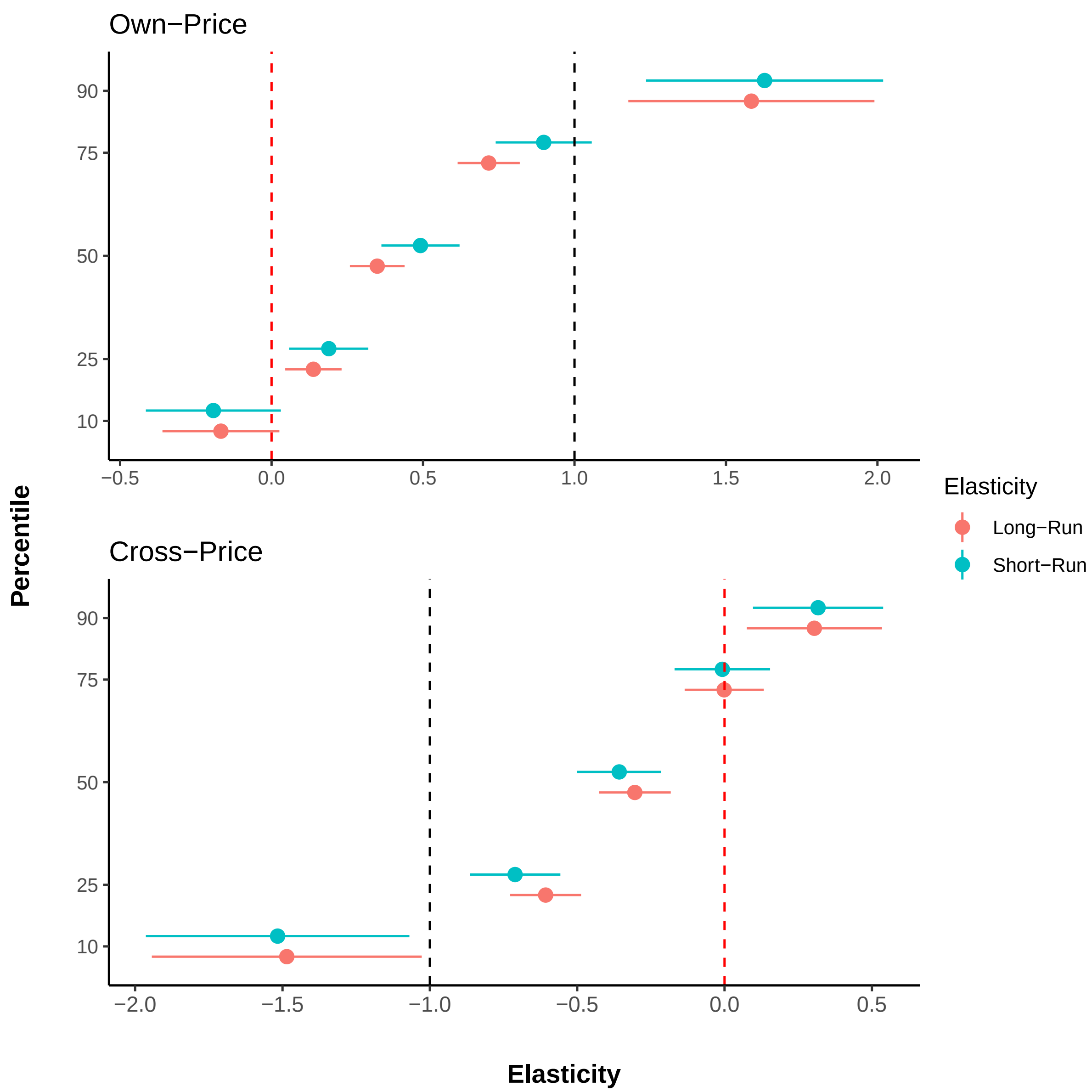


Figure 1: Commodity Price Locations Continuously Observed Between 2004 and 2016



Note: Point-and-whisker plots show the mean of the statistic's percentile using at the dots and the 95% confidence interval using the whiskers. Plots in red illustrate the bootstrapped results for the Corn to Corn (ϕ^{CC}) *equation and the green plots illustrate the bootstrapped results for the Other to Corn* (ϕ^{OC}) *equation. For reference, zero was indicated with a dashed red line.*

Figure 2: Summary of Markov Transition Regression Coefficients Across Regions



Note: Point-and-whisker plots show the mean of the statistic's percentile using at the dots and the 95% confidence interval using the whiskers. Plots in red illustrate the bootstrapped results for the long-run elasticity terms and the green plots illustrate the bootstrapped results for the short-run elasticity terms. For reference, zero was indicated with a dashed red line and a black dashed line was placed at 1 or -1 when appropriate.

Figure 3: Summary of Short-run and Long-run Corn Planting Elasticities Across Regions

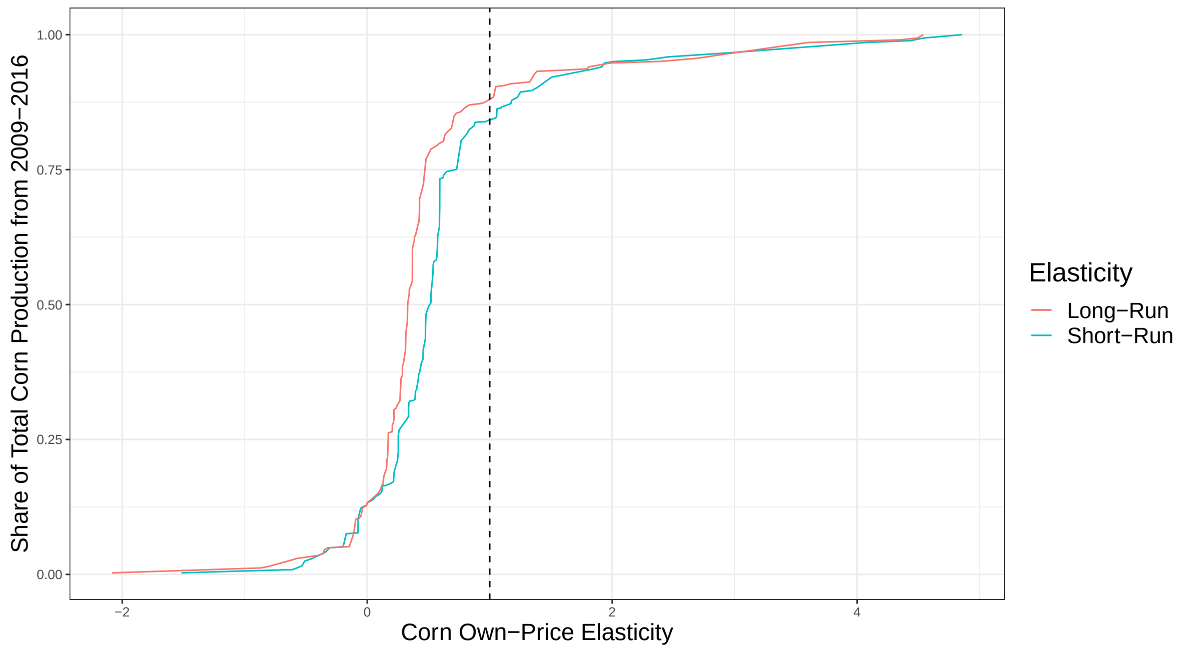


Figure 4: Cumulative Share of Corn Production by Own-Price Corn Planting Elasticity

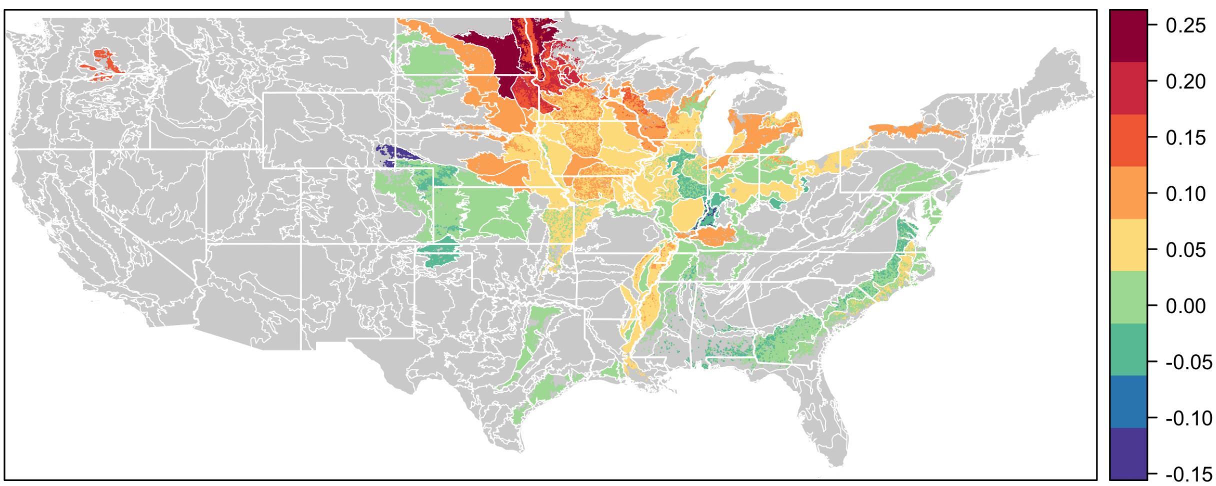


Figure 5: Average MLRA-Group Short-Run Own-Price Planting Marginal Effect over US

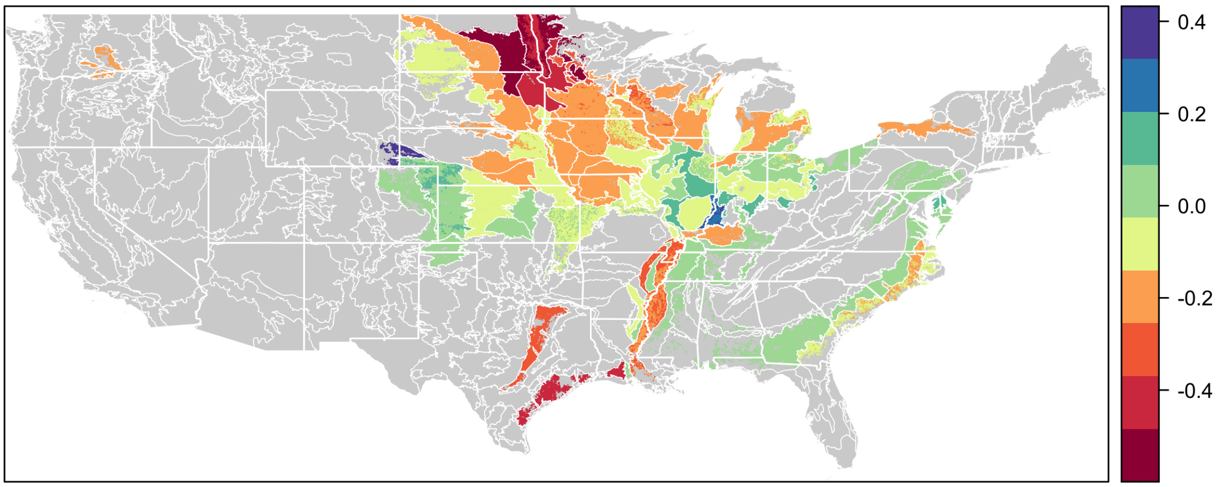


Figure 6: Average MLRA-Group Short-Run Cross-Price Planting Marginal Effect over US

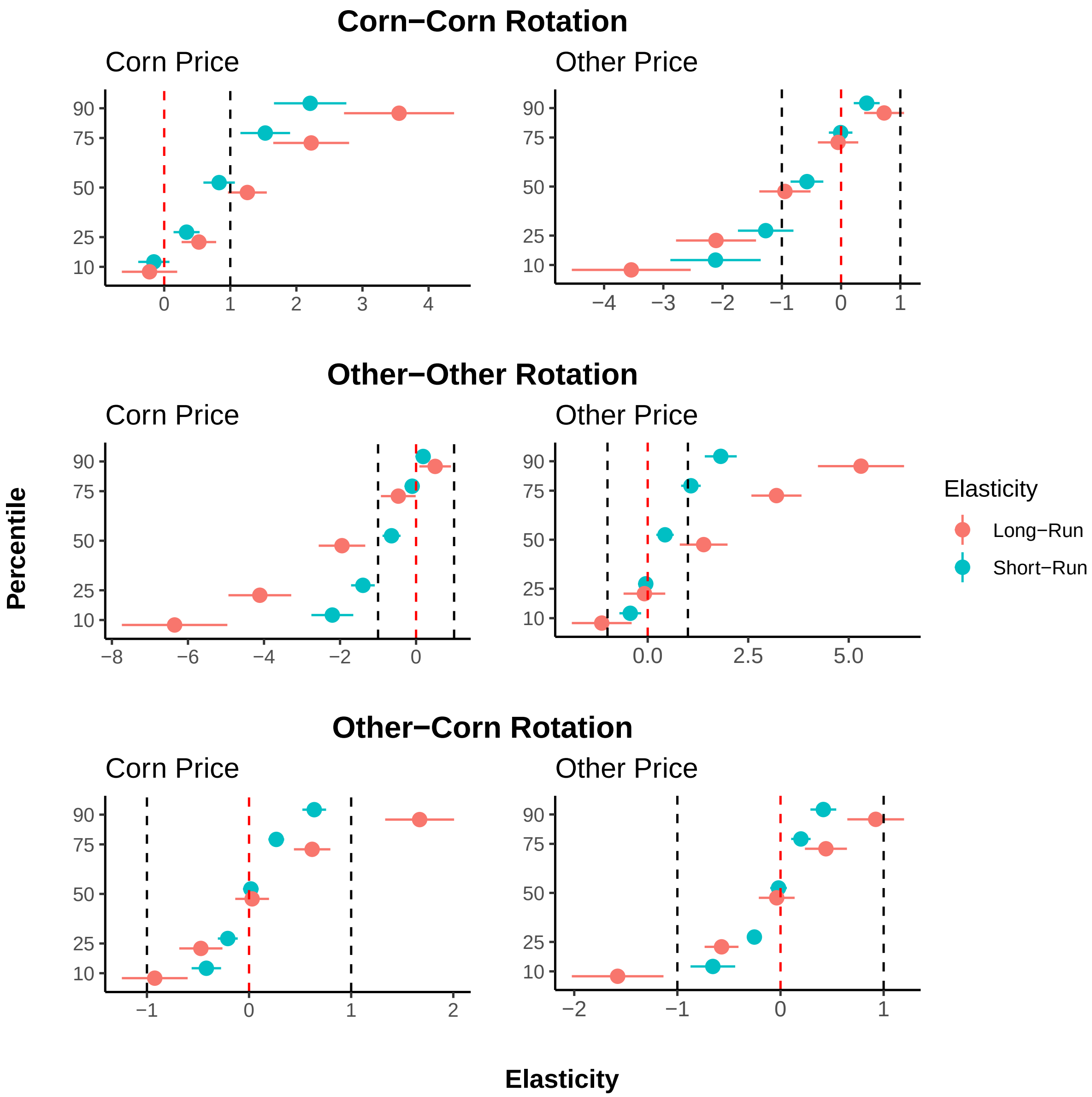


Figure 7: Short- and Long-Run Rotation Elasticity Summaries Across Regions

Note: Point-and-whisker plots show the mean of the statistic's percentile using at the dots and the 95% confidence interval using the whiskers. Plots in red illustrate the bootstrapped results for the long-run elasticity terms and the green plots illustrate the bootstrapped results for the short-run elasticity terms. For reference, zero was indicated with a dashed red line and a black dashed line was placed at 1 or -1 when appropriate.

Tables

Table 1: Data Summary

Statistic	Mean	Std. Deviation
Corn Plant Prob.	41.65%	49.30%
Field Size (acres)	58.76	63.71
Irrigation Status	8.52%	27.92%
Field Slope	3.14	3.49
NCCPI	0.57	0.21
Historic Dry Year	43.75%	49.61%
Historic Wet Year	48.64%	49.98%
Corn Yield (bu)	149.45	28.84
Number of Observations	30,124,818	
Number of MLRAs	68	
Number of MLRA-groups	115	
Total Field Acreage	220,858,765	

Table 2: Observed Share of NASS Corn Acreage by Year

Year	NASS Reported Acreage	Total Observed Acreage	Obs. Share of NASS Value
2009	86,382,000	70,877,484	82.05%
2010	88,192,000	72,831,712	82.58%
2011	91,936,000	75,895,029	82.55%
2012	97,291,000	80,450,263	82.69%
2013	95,365,000	79,642,717	83.51%
2014	90,597,000	74,352,013	82.07%
2015	88,019,000	73,577,582	83.59%
2016	94,004,000	80,357,993	85.48%
Total	731,786,000	607,984,794	83.08%

Table 3: Nationwide Corn Elasticities With and Without Heterogeneity

Model →	National				Corn Belt	
	Het. Aggr.		Pooled		Het. Aggr.	
Statistic Weighting →	Acreage	Quantity	Acreage	Quantity	Acreage	Quantity
Own-Price SR Elasticity	0.715***	0.687***	0.529***	0.503***	0.386***	0.383***
	(0.067)	(0.068)	(0.019)	(0.018)	(0.058)	(0.058)
Own-Price LR Elasticity	0.574***	0.541***	0.454***	0.432***	0.253***	0.252***
	(0.045)	(0.045)	(0.016)	(0.015)	(0.036)	(0.036)
Own-Price Lag Dep.	-0.247	-0.269	-0.164	-0.165	-0.525	-0.517
Cross-Price SR Elasticity	-0.549***	-0.515***	-0.373***	-0.355***	-0.227***	-0.229***
	(0.071)	(0.069)	(0.015)	(0.015)	(0.062)	(0.062)
Cross-Price LR Elasticity	-0.467***	-0.422***	-0.314***	-0.298***	-0.146***	-0.149***
	(0.062)	(0.056)	(0.013)	(0.012)	(0.041)	(0.04)
Cross-Price Lag Dep.	-0.175	-0.221	-0.189	-0.191	-0.562	-0.543

Note: Standard errors shown in parentheses. Asterisks *** ,** , and * indicate significance level at 1%, 5%, and 10% respectfully. The lag dependent variable refers to the relative difference between the long- and short-run elasticities $\left(\frac{\varepsilon_{LR}-\varepsilon_{SR}}{\varepsilon_{LR}}\right)$, which is analogous to the coefficient on the lagged dependent variable in a single equation dynamic model.

Table 4: Nationwide Mean Elasticity Differences Between Heterogeneous and Pooled Models ($\varepsilon_{Het} - \varepsilon_{Pooled}$)

Elasticity	Crop: Corn	Crop: Other
SR Acre-Weighted	0.165***	−0.169***
	(0.069)	(0.072)
LR Acre-Weighted	0.118***	−0.154***
	(0.048)	(0.063)
Lagged Dep. Acre-Weighted	−0.036	0.037
	(0.028)	(0.042)
SR Qty.-Weighted	0.163***	−0.153**
	(0.07)	(0.07)
LR Qty.-Weighted	0.109***	−0.124**
	(0.047)	(0.057)
Lagged Dep. Qty.-Weighted	−0.058**	−0.009
	(0.028)	(0.041)

Note: Standard errors shown in parentheses. Asterisks *** ,** , and * indicate significance level at 1%, 5%, and 10% respectfully. The lag dependent variable refers to the relative difference between the long- and short-run elasticities $\left(\frac{\varepsilon_{LR}-\varepsilon_{SR}}{\varepsilon_{LR}}\right)$, which is analogous to the coefficient on the lagged dependent variable in a single equation dynamic model.

Table 5: Transitional and Sequential Probabilities In and Outside of the Corn Belt

Region	Probability	Crop Choices			
		Corn→Corn	Other→Corn	Corn→Other	Other→Other
Outside Corn Belt	Sequential	0.128	0.289	0.289	0.294
	Transitional	0.307	0.495	0.693	0.505
Corn Belt	Sequential	0.179	0.369	0.369	0.082
	Transitional	0.327	0.817	0.673	0.183

Supplementary Appendix

Source of Pooling Bias in Autoregressive Cropping Sequences

To describe how pooling bias arises in crop-sequence estimation, we follow Pesaran and Smith (1995) and suppose that our relationship of interest exhibits a lagged dependent structure such as the one in equation A1. Here y_{it} is our outcome of interest with i indexing individuals and t indexing time. In our application, y_{it} stands for cropping choices. We allow for unobserved heterogeneity at the individual level through α_i and the covariate of interest is x_{it}. Here v_{it} is our modeled error exhibited by equation A2 which shows that we have unmodeled heterogeneity in our lagged dependent variable coefficient as well and in the coefficient on the covariate. Here ε_{it} is i.i.d.

(A1) $y_{it} = \alpha_i + \lambda y_{it-1} + \beta' x_{it} + v_{it}$

(A2) $v_{it} = \varepsilon_{it} + \eta_{1i} y_{it-1} + \eta_{2i}' x_{it}$

For simplicity, suppose that $E[x_{it}] = 0$ for all individuals and time periods. Suppose there exists autocorrelation in x_{it} such that $\gamma_i(s) = \mathbb{E}[x_{it} x_{it+s}] = \mathbb{E}[x_{it} x_{it-s}]$. In this case, x_{it} will be correlated with the error term according to equation A3 (Pesaran and Smith, 1995). Here λ_i is the true underlying individual specific lagged coefficient where $\lambda_i = \lambda + \eta_{1i}$ and β_i is the true heterogeneous coefficient of the other covariate such that $\beta_i = \beta + \eta_{2i}$. Note that x_{it} is uncorrelated with the error term only under certain conditions. First, if there is no autocorrelation in x_{it}, then $\gamma_i(s)$ is zero for every value of s. Second, if there is no heterogeneity in the lagged dependence structure (i.e., if $\eta_{1i} = 0$ which implies $\lambda_i = \lambda$).

(A3) $$\mathbb{E}[x_{it}v_{it}] = \sum_{j=0}^{\infty} \mathbb{E}\left[\eta_{1i}\beta_i\lambda_i^j\right]\gamma_i(|j+1|)$$

Next consider the correlation of the lagged dependent variable with the error term. The correlation between the lagged dependent variable and the error term can be stated as equation A4 (Pesaran and Smith, 1995). The first term in equation A4 arises due to correlation between the x terms that exist in the lagged dependent values and the other higher order lagged values that are embedded in the functional form of each of the lagged dependent values. The many terms here arise from the fact that every lagged value of x_{it} is present in the y_{it-1} term left unmodeled in the heterogeneous lagged coefficient and in the y_{it-1} term itself. The second term arises due to direct correlation between contemporaneous lagged dependent variable (y) terms embedded inside of the y_{it-1} term. The σ_i^2 term represents the variance of the i.i.d. y_{it} term. Notice that the second term of equation A4 is not equal to zero even if there is no autocorrelation in x_{it}. The final term is the correlation between the lagged values of x_{it} embedded in y_{it-1} and the unmodeled individual-specific effect from x_{it} in the error term.

(A4)
$$\mathbb{E}[y_{it-1}v_{it}] = \sum_{s=0}^{\infty}\sum_{r=0}^{\infty}\left\{\mathbb{E}\left[\eta_{1i}\beta_i^2\lambda_i^{r+s}\right]\gamma_i(|r-s|)\right\} + \sigma_i^2\sum_{s=1}^{\infty}\left\{\mathbb{E}\left[\eta_{1i}\lambda_i^{2s}\right]\right\} + \sum_{s=0}^{\infty}\left\{\mathbb{E}\left[\eta_{2i}\beta_i\lambda_i^s\right]\right\}\gamma_i(|s+1|)$$

It is clear from equations A3 and A4 that pooling bias will persist so long as we have unmodeled heterogeneity in the lagged dependent structure in the model ($\eta_{1i} \neq 0$). Further bias can exist if the regressor is autocorrelated. It is therefore important to incorporate as much heterogeneity in the lagged dependent relationship as possible to ensure that the effects of prices on rotations is accurately modeled.

Detailed Explanation of a Markov Chain

The structure of a Markov chain supposes that the data follows a Markov process of a certain order. A first order Markov process supposes that the cropping decisions are a function of contemporaneous conditions and the prior land cover. If this is the case then we can write the probability of selecting corn in a given year as a function of probabilities of transitioning to corn given the prior crop choice.

For complete understanding of how Markov chains are estimated, it is important to go over definitions of probability terms. The joint probability of two events X and Y occurring is written as equation A5. In our case, we suppose that crop choice follows a first order Markov process. In this case, the probability of planting corn this year (event X) is conditional on whether corn was planted last year (event Y). With this simple relabeling, we can construct the probability of corn being planted in two consecutive years as equation A6. We define our variable of interest y_{it} as in equation A7.

(A5) $Pr\left[X \cap Y\right] = Pr\left[X \mid Y\right] Pr\left[Y\right]$

(A6) $Pr\left[y_{it} = 1 \cap y_{it-1} = 1\right] = Pr\left[y_{it} = 1 \mid y_{it-1} = 1\right] Pr\left[y_{it-1} = 1\right]$

(A7) $y_{it} = \begin{cases} 1 \mid & \text{Corn planted on field } i \text{ at time } t \\ 0 \mid & \text{Some other crop is planted} \end{cases}$

To estimate the unconditional probability we need only exhaust the conditions. We consider only observations that were planted to corn or one of the corn alternatives in two consecutive years. That is, $y_{it} = 1$ or $y_{it} = 0$ for all i and t in our observation sets. That is, we do not consider fields that did not have some non-crop cover in the prior period. In this way we exclusively model crop choice at the intensive margin as opposed to the extensive margin. From here, we can estimate the *unconditional* probability of planting corn by including the additional term in equation A8. The unconditional probability will then be the sum of the joint probability states shown in equation A9.

(A8) $$Pr\left[y_{it} = 1 \cap y_{it-1} = 0\right] = Pr\left[y_{it} = 1 \mid y_{it-1} = 0\right]\left(1 - Pr\left[y_{it-1} = 1\right]\right)$$

(A9) $$Pr\left[y_{it} = 1\right] = Pr\left[y_{it} = 1 \cap y_{it-1} = 1\right] + Pr\left[y_{it} = 1 \cap y_{it-1} = 0\right]$$

$$= Pr\left[y_{it} = 1 \mid y_{it-1} = 1\right] Pr\left[y_{it-1} = 1\right] + Pr\left[y_{it} = 1 \mid y_{it-1} = 0\right]\left(1 - Pr\left[y_{it-1} = 1\right]\right)$$

Here we have the unconditional probability of planting corn in a given year t as a function of the *lagged* probability of planting corn. To simplify notation, we use $\mathbb{P}_{it}$ as the unconditional probability of planting corn for field i at time t.

(A10) $$\mathbb{P}_{it} = \mathbb{P}_{it-1} Pr\left[y_{it} = 1 \mid y_{it-1} = 1\right] + \left(1 - \mathbb{P}_{it-1}\right) Pr\left[y_{it} = 1 \mid y_{it-1} = 0\right]$$

This form highlights that the probability of planting corn depends on two transition probabilities—the probability of planting corn given corn was previously planted and the probability of planting corn given that an other crop was previously planted. These transition probabilities depend on the relative profitability of planting the same crop type twice in consecutive periods versus alternating between crops. The incentives to rotate crops are discussed in detail in Hennessy (2006) but we will discuss them here as well for completeness. The per-acre profit of joint inter-periodic planting decisions are laid out in equations A11, A12, A13, and A14 below. By alternating crops each year, the producer gains some positive "boost" to yields. The boost to corn yields for field i at time t is B_{it}^{OC}, and the boost to corn-alternative yields is B_{it}^{CO}.

(A11) $\pi\left(y_{it}=1 \mid y_{it-1}=1\right)=\pi_{it}^{CC}$

(A12) $\pi\left(y_{it}=1 \mid y_{it-1}=0\right)=\pi_{it}^{CC}+B_{it}^{OC} P_{it}^{C}$

(A13) $\pi\left(y_{it}=0 \mid y_{it-1}=1\right)=\pi_{it}^{OO}+B_{it}^{CO} P_{it}^{O}$

(A14) $\pi\left(y_{it}=0 \mid y_{it-1}=0\right)=\pi_{it}^{OO}$

If the crop transition probability is a function of the differences between conditional profits, it is clear that the transition probabilities will be a function of the yield improvements from rotations and the contemporary prices of each of the crops. We show this in equations A15 and A16. Here the functions $g\left(\cdot\right)$ and $h\left(\cdot\right)$ are non-decreasing confining to the properties of probability (e.g. they are positive but less than one).

(A15)

$$Pr\left[y_{it}=1 \mid y_{it-1}=0\right]=g\left(\pi\left(y_{it}=1\right)-\pi\left(y_{it}=0\right) \mid y_{it-1}=0\right)=g\left(\pi_{it}^{CC}-\pi_{it}^{OO}+B_{it}^{OC}P_{it}^{C}\right)$$

(A16)

$$Pr\left[y_{it}=1 \mid y_{it-1}=1\right]=h\left(\pi\left(y_{it}=1\right)-\pi\left(y_{it}=0\right) \mid y_{it-1}=1\right)=h\left(\pi_{it}^{CC}-\pi_{it}^{OO}-B_{it}^{CO}P_{it}^{O}\right)$$

The conditions that define profit maximizing crop choices are written in equation A17. We estimate one model using data where the lagged crop choice was some corn alternative ($y_{it-1}=0$). We estimate a second model using only observations where corn was the lagged crop choice ($y_{it-1}=1$) to estimate the corn-to-corn transition models. We then use a dummy variable for the contemporary crop choice y_{it} equal to one if corn were selected in the contemporary period and zero otherwise. In this way, we directly estimate the conditional transition probabilities needed in equation A10.

$$\text{(A17)}\quad y_{it}^{\star}=\begin{cases} 1 \mid \pi_{it}^{CC}-\pi_{it}^{OO}+B_{it}^{OC}P_{it}^{C} \geqslant \varepsilon_{it}^{OC} \text{ with } y_{it-1}=0 \\ 0 \mid \pi_{it}^{CC}-\pi_{it}^{OO}+B_{it}^{OC}P_{it}^{C} < \varepsilon_{it}^{OC} \text{ with } y_{it-1}=0 \\ 1 \mid \pi_{it}^{CC}-\pi_{it}^{OO}-B_{it}^{CO}P_{it}^{O} \geqslant \varepsilon_{it}^{CC} \text{ with } y_{it-1}=1 \\ 0 \mid \pi_{it}^{CC}-\pi_{it}^{OO}-B_{it}^{CO}P_{it}^{O} < \varepsilon_{it}^{CC} \text{ with } y_{it-1}=1 \end{cases}$$

If the epsilon terms in equation A17 follow an extreme value distribution, the logistic regression is valid for estimation. However, suppose for the moment that we estimate our conditional transition probabilities using a set of two linear probability models according to equations A18 and A19. Here our variables X_{it} are controls, common to both equations and the parameters with a zero subscript are scalar intercept terms. Combining these two identities in equation A10 gives us equation A20. After simplifying, this gives equation A21. This shows that, if approximated as a linear probability model, our base model can be represented as a first-order autoregressive model with additional interactions between the lagged independent variable and the controls.

$$\text{(A18)}\quad [y_{it} = 1 \mid y_{it-1} = 0] = \beta_0^0 + X_{it}\beta^0 + \varepsilon_{it}^0$$

$$\text{(A19)}\quad [y_{it} = 1 \mid y_{it-1} = 1] = \beta_0^1 + X_{it}\beta^1 + \varepsilon_{it}^1$$

$$\text{(A20)}\quad y_{it} = y_{it-1}\left(\beta_0^1 + X_{it}\beta^1\right) + (1 - y_{it-1})\left(\beta_0^0 + X_{it}\beta^0\right)$$

$$\text{(A21)}\quad y_{it} = \beta_0^0 + X_{it}\beta^0 + y_{it-1}\left[\beta_0^1 - \beta_0^0\right] + y_{it-1}\left[X_{it}\left(\beta^1 - \beta^0\right)\right]$$

Our interest in estimating the transition probabilities is to calculate the steady state probability of planting corn. This is akin to an equilibrium condition in the sense that this is the probability that producers will plant corn given the soil, weather, and marketing environment remain constant. We functionally represent this by setting the unconditional probability of growing corn in the lagged periods equal to those in the contemporaneous period in equation A10 as shown in A22. After rearranging, we get equation A23.

$$\text{(A22)}\quad \mathbb{P}_i = \mathbb{P}_i Pr\left[y_{it} = 1 \mid y_{it-1} = 1\right] + \left(1 - \mathbb{P}_i\right) Pr\left[y_{it} = 1 \mid y_{it-1} = 0\right]$$

$$\text{(A23)}\quad \mathbb{P}_i = \frac{Pr\left[y_{it} = 1 \mid y_{it-1} = 0\right]}{\left[1 - Pr\left[y_{it} = 1 \mid y_{it-1} = 1\right] + Pr\left[y_{it} = 1 \mid y_{it-1} = 0\right]\right]}.$$

From here, we utilize equations A22 and A23 to construct our price effects. We create distinct effects by assuming different components of equation A10 are subject to change. To capture the short-run effects, we hold constant the steady state probabilities (i.e., $\frac{\partial \mathbb{P}_i}{\partial X} = 0$). In this way, the short-run effect represents the effect of a price shock that only affects the transition probabilities and does not impact the lagged crop choice. To capture the long-run effect, we take the derivative of equation A23. The long-run effect represents the change in the probability of planting corn after allowing the probability of corn to reach a new steady-state.

A Discussion on the Long and Short-Run Effects

Here we further describe the reasons for the dynamics we observe in crop choices in the long run and short run. We suppose that there is some lagged dependence between our crop choices such as those from rotation incentives. Specifically, we suppose that there are benefits to alternating between crops in each year. In this case, the crop choice process should exhibit negative dependence. Such as equation A24 show below. This is the empirical model in a linear probability model.

$$\text{(A24)}\quad y_{it} = \beta_0 + y_{it-1}\gamma + X_{it}\beta + y_{it-1}X_{it}\rho + \varepsilon_{it}$$

In our models we estimate the effect of a change in the steady state price of a crop. That is, the effect of a persistent price shock. We now consider how a general persistent shock starting in period ($t^\star$) impacts a negatively dependent sequence with all else being equal. We can express this hypothetical series with one that does not experience the shock. This shock $s^\star_{it} = s > 0$ for all $t \geqslant t^\star$ and zero otherwise.

$$\text{(A25)}\quad y^\star_{it} = \beta_0 + y_{it-1}\gamma + X_{it}\beta + y_{it-1}X_{it}\rho + \varepsilon_{it} + s^\star_{it}$$

The short-run effect from the shock in the initial period is readily derived since the lagged values of both series should be the same.

(A26)

$$y^\star_{it^\star} - y_{it^\star} = \left[\cancel{\beta_0} + y^\star_{it^\star-1}\gamma + \cancel{X_{it^\star}\beta} + y^\star_{it^\star-1}X_{it^\star}\rho + \cancel{\varepsilon_{it^\star}} + s\right] - \left[\cancel{\beta_0} + y_{it^\star-1}\gamma + \cancel{X_{it^\star}\beta} + y_{it^\star-1}X_{it^\star}\rho + \cancel{\varepsilon_{it^\star}}\right]$$

$$= \left(y^\star_{it^\star-1} - y_{it^\star-1}\right)\gamma + \left(y^\star_{it^\star-1} - y_{it^\star-1}\right)X_{it^\star}\rho + s = s$$

The effect of a permanent shock is captured as where s is always a trailing figure in the difference.

$$\text{(A27)}\quad y^{\star}_{it^{\star}} - y_{it^{\star}} = \left(y^{\star}_{it^{\star}-1} - y_{it^{\star}-1}\right)\gamma + \left(y^{\star}_{it^{\star}-1} - y_{it^{\star}-1}\right) X_{it^{\star}}\rho + s$$

In this case our next difference is:

$$\text{(A28)}\quad y^{\star}_{it+1^{\star}} - y_{it+1^{\star}} = \left(y^{\star}_{it^{\star}} - y_{it^{\star}}\right)\gamma + \left(y^{\star}_{it^{\star}} - y_{it^{\star}}\right) X_{it^{\star}+1}\rho + s = \left[1 + \left(\gamma + X_{it^{\star}+1}\rho\right)\right] s$$

For notational ease, consider our X values do not vary over time. We can then express the general sequence out to the k^{th} period in equation A29.

$$\text{(A29)}\quad y^{\star}_{it^{\star}+k} - y_{it^{\star}+k} = \left[1 + \sum_{j=1}^{k} \left(\gamma + X_i\rho\right)^j\right] s$$

Note that the $1+\sum_{j=1}^{k} (\gamma + X_i\rho)^j$ term is a geometric series. Again, if $|\gamma + X_i\rho| < 1$, then this series converges to $\frac{1}{1-\gamma-X_i\rho}$. Therefore, the entire sequence will converge to equation A30.

$$\text{(A30)}\quad \lim_{k\to\infty} \left(y^{\star}_{it+k^{\star}} - y_{it+k^{\star}}\right) = \left[\frac{1}{1-\gamma-X_i\rho}\right] s$$

Under negative dependence (that is, if $\gamma + X_i\rho < 0$), this will converge to a long-run effect *below* the short-run effect. This is likely to hold when there are economic benefits to alternating between crops over time.

Intuitively, the longer term benefits of rotating crops will, to some degree, cause producers to revert back to technically favorable rotations over time. However, these rotations may be worth abandoning temporarily due to a favorable idiosyncratic price shocks. Note that, by the way we characterized s, the effect can arise from a permanent shock (e.g. an intercept shift), or a shock to our regressor values such as a permanent shift in prices. Figure A1 provides an numerical illustration of short-run and long-run effects from a permanent shock to a simple negative dependent autoregressive sequence. The initial, short-run jump is larger than the long-run change. This suggests that negatively dependent cropping choices imply that producers' planting response will be larger in the short-run than in the long-run.

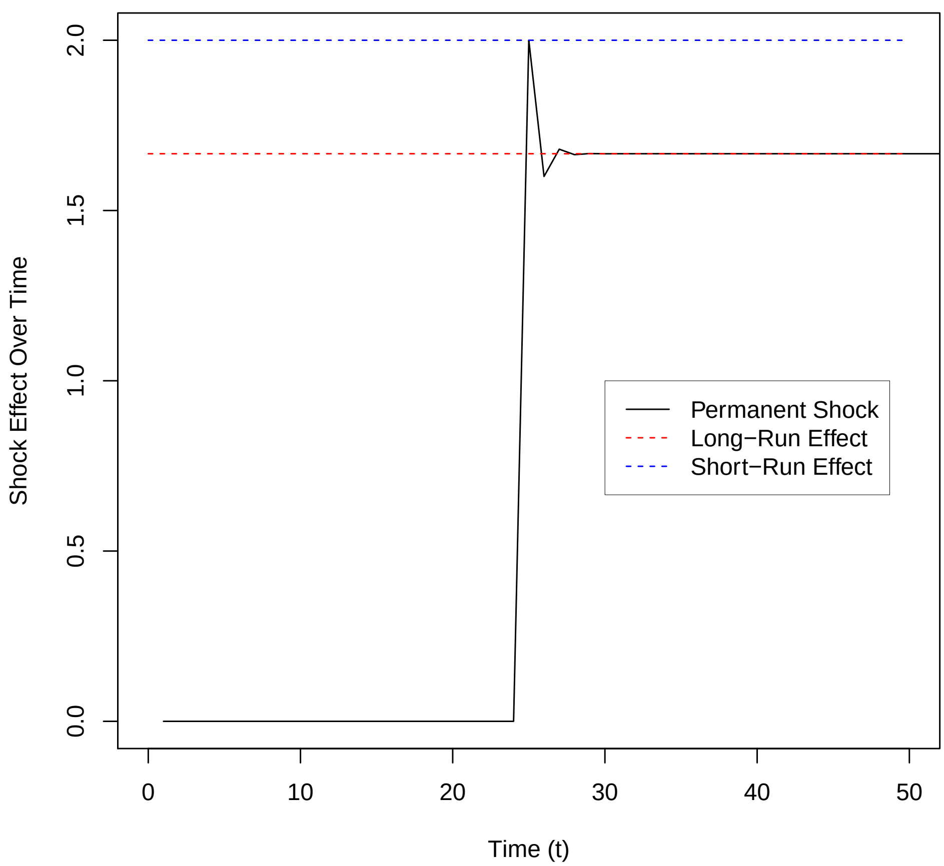


Figure A1: Short- and Long-Run Effect Illustration Under Negative Dependence

Note: The black line represents the autoregressive sequence. To construct the short and long-run effects of a shock we apply a permanent shock in period $t = 25$ of size 2 which persists in every period after period 25. The sequences exhibit negative dependence with a lagged coefficient value of -0.2.

Cropland Data Layer Classification Accuracy Rates by Crop

Here we present a summary of the Cropland Data Layer 2016 classification accuracy rates across the states in the analysis. We do this to give an idea of how accurate corn classification rates are relative to some of the more prominent corn alternatives. Tables A1 and A2 show statistics on the producer accuracy rates and user accuracy rates. The producer accuracy rate is the probability that the classification matches the class for given ground-truthed observations. The user accuracy rate is the probability that the ground-truthed class matches the classification estimate. Formally, the producers accuracy rate is $\mathbb{P}\left[c \mid c^{\star}\right]$ and the user accuracy rate is $\mathbb{P}\left[c^{\star} \mid c\right]$. Here (c) indicates a classification of crop "c" and $(c^{\star})$ indicates the field was ground-truthed to crop "c". A high producer accuracy rate relative to user accuracy rate is a symptom of over-classifying and a high user accuracy rate relative to the producer accuracy rate could be a symptom of under-classifying. For instance, if a classifier indiscriminately classified an entire county as corn, its producer accuracy would be 100% but its user accuracy rate would be the probability that the pixel were ground-truthed to crop "c". If, on the other hand, only classified a single ground-truthed corn pixel were classified to corn, it would have a 100% user accuracy rate. However, it may fail to correctly classify the vast majority of the ground-truthed corn pixels in the state.

Tables A1 and A2 show that corn is classified more accurately relative to many of its alternatives. The majority of states have accuracy rates in the mid to high 90s for both the user and producer accuracy rates. This suggests that the CDL is good at avoiding over- and under-classifying corn pixels. Soybeans, cotton and rice generally have good accuracy rates but they are less consistently accurate across the states in our analysis than corn pixels. However, the classifier performs significantly worse on wheat varieties, fallowing, and double-cropping land covers when compared to corn. These classification issues may be due, in part, to similarities in spectral characteristics between different corn alternatives. We help mitigate these classification error rates by combining corn alternatives into a single "other" category.

Land Cover	Min	1st Quartile	Median	Mean	3rd Quartile	Max
Corn	76.28	93.17	94.95	94.10	97.23	99.10
Cotton	47.61	86.54	92.09	87.38	93.34	94.37
Soybeans	26.81	87.53	94.12	85.95	97.16	98.02
Dbl Crop WinWht/Cotton	21.47	37.84	46.65	51.54	68.18	87.76
Rice	15.35	73.46	92.99	77.18	96.72	97.58
Spring Wheat	10.11	29.98	52.41	57.43	89.92	97.26
Fallow/Idle Cropland	8.57	28.84	47.70	51.45	71.25	97.42
Winter Wheat	8.37	40.90	72.88	66.31	94.97	98.41
Canola	6.25	26.61	72.10	56.21	77.24	96.69
Dbl Crop Barley/Soybeans	0	52.48	60.33	52.62	62.04	76.70
Dbl Crop WinWht/Soybeans	0	35.80	76.80	64.20	86.97	94.83
Buckwheat	0	29.66	47.32	48.26	68.75	94.74
Durum Wheat	0	25.06	38.10	47.58	72.41	100
Sorghum	0	23.38	40.34	39.37	47.70	91.82
Rye	0	17.54	32.56	34.49	51.27	100
Dbl Crop Barley/Corn	0	16.75	34.79	36.00	34.97	93.75
Dbl Crop WinWht/Corn	0	12.22	27.66	29.36	45.79	76.07
Dbl Crop WinWht/Sorghum	0	12.17	33.08	28.15	37.90	73.80
Barley	0	9.46	22.72	33.06	58.62	87.12
Dbl Crop Soybeans/Oats	0	7.55	25.35	28.62	48.59	82.52
Dbl Crop Oats/Corn	0	5.14	24.39	26.65	33.49	100
Dbl Crop Corn/Soybeans	0	4.15	8.31	10.45	14.86	26.26

Table A1: Cropland Data Layer 2016 Producer Accuracy Rate Statistics

Land Cover	Min	1st Quartile	Median	Mean	3rd Quartile	Max
Rice	91.74	95.90	97.65	96.59	98.27	98.47
Corn	82.72	89.96	94.94	92.94	96.79	98.82
Cotton	79.19	86.52	91.61	90.14	94.78	97.07
Soybeans	64.84	85.67	93.15	89.89	95.76	98.60
Dbl Crop WinWht/Cotton	52.89	75.28	80.11	80.16	90.27	98.10
Spring Wheat	36.00	70.62	76.89	76.63	89.47	100.00
Winter Wheat	31.90	68.32	78.45	77.75	92.78	96.87
Fallow/Idle Cropland	22.43	56.48	72.88	71.01	82.86	98.05
Canola	16.67	84.97	93.55	80.43	97.61	100.00
Dbl Crop Barley/Soybeans	0.00	79.88	80.72	74.62	87.45	98.35
Durum Wheat	0.00	78.89	80.18	71.44	82.30	97.56
Dbl Crop WinWht/Soybeans	0.00	77.89	84.23	75.85	87.12	100.00
Sorghum	0.00	64.13	77.62	68.19	86.99	96.33
Dbl Crop Barley/Corn	0.00	61.81	70.12	63.46	81.25	100.00
Dbl Crop WinWht/Corn	0.00	60.61	70.21	62.77	76.00	97.37
Dbl Crop Soybeans/Oats	0.00	55.67	73.10	60.22	85.58	100.00
Barley	0.00	50.48	63.37	60.58	81.24	95.95
Rye	0.00	42.86	59.57	59.53	77.40	100.00
Dbl Crop WinWht/Sorghum	0.00	39.80	61.52	57.12	78.95	96.39
Dbl Crop Corn/Soybeans	0.00	31.25	52.28	51.00	70.99	100.00
Dbl Crop Oats/Corn	0.00	30.36	62.75	56.48	91.85	100.00
Buckwheat	0.00	28.87	75.00	60.30	87.77	100.00

Table A2: Cropland Data Layer 2016 User Accuracy Rate Statistics

Intratemporal Price Correlation

Figures A2 and A3 show the box plots for the corn price and other price index for each year of the analysis. These plots show evidence of consistent basis patterns arising over time since the variance is approximately identical in each year. They also show that in many years, the mean price observation escapes the price distributions in subsequent years. This illustrates that the price variation is larger between years than it is within years. Taken together this means there are likely within-year dependence problems in the dataset.

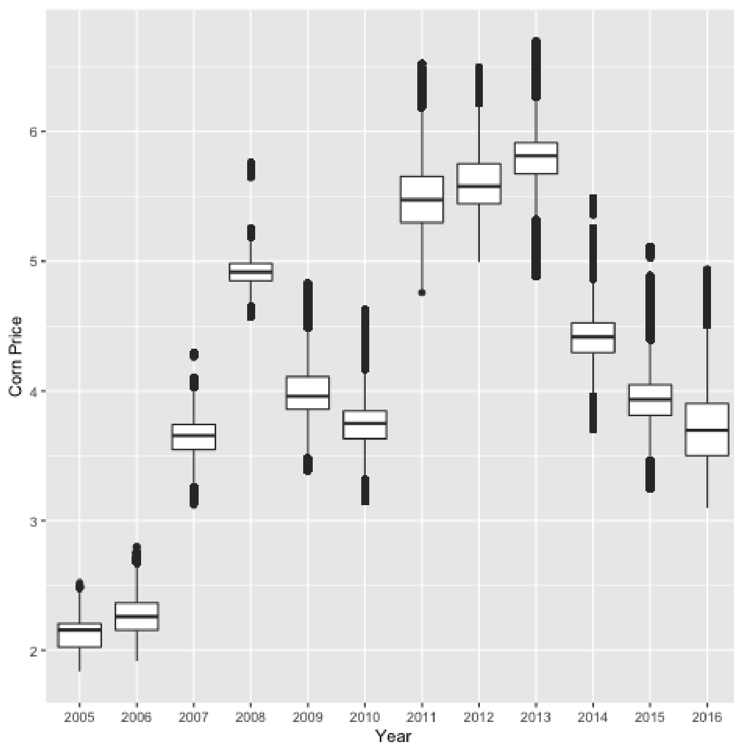


Figure A2: Corn Price Distributions by Year

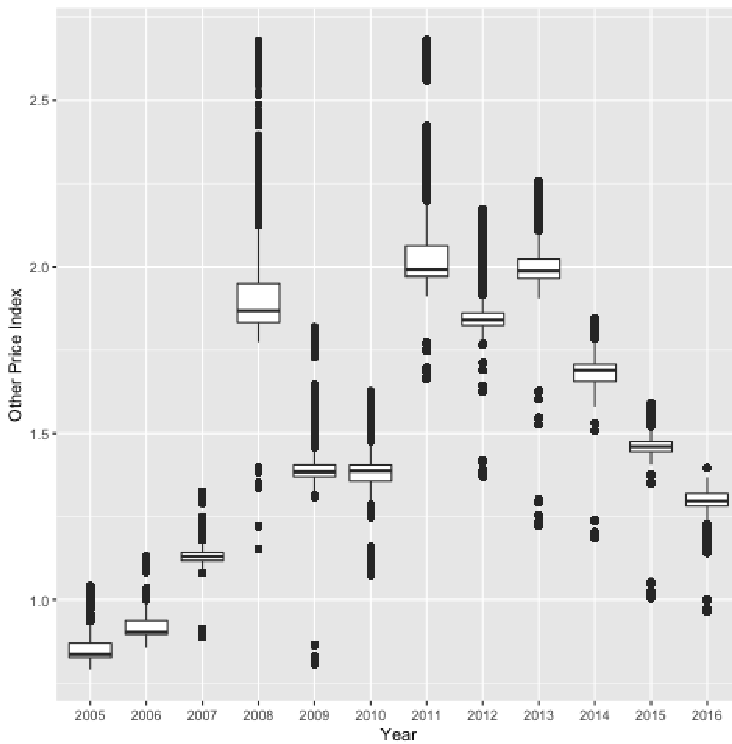


Figure A3: Other Price Distributions by Year

Minimum Year by MLRA

Figure A4 shows the MLRA map and earliest year of analysis for each MLRA dataset.

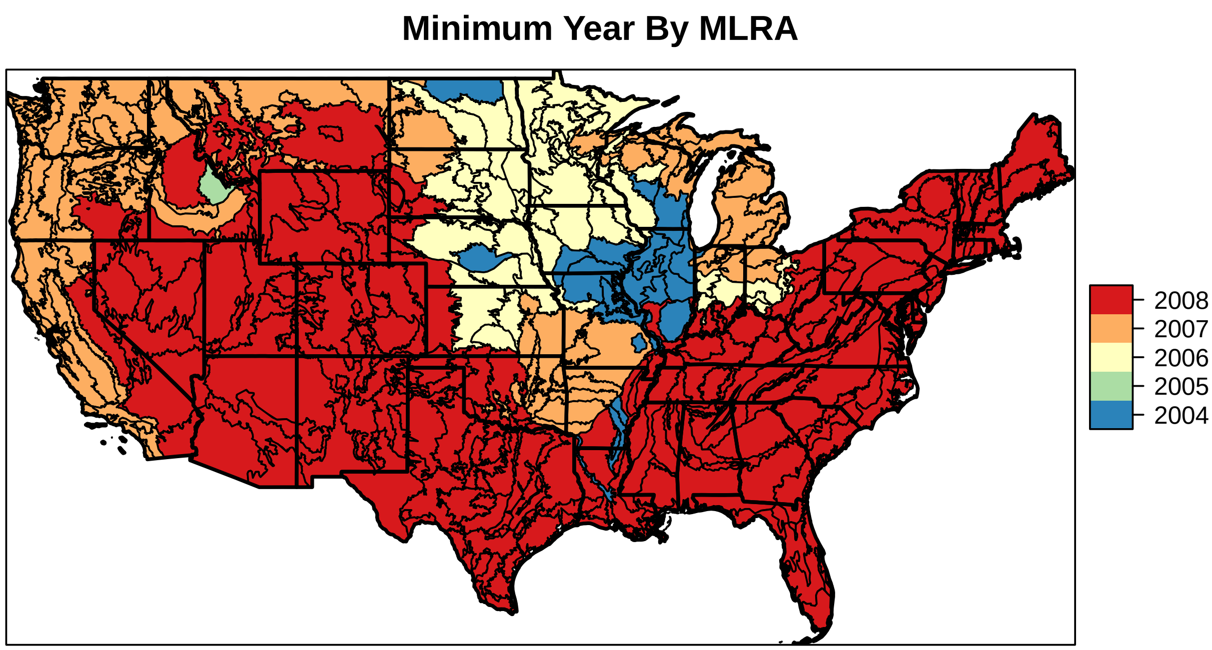


Figure A4: Minimum Observation Year by Major Land Resource Area

Combining Soil Texture Groups by ”Closeness”

This study estimated separate models using MLRAs and soil texture. Soil texture groups are defined based off of the soil's composition of silt, sand, and clay. Traditionally these are aggregated to 12 groups but due to data limitations, we use the 5-group designation in table A3. In some instances we needed to further combine observations by soil texture so that enough observations to robustly estimate each of the Markov transition regressions. We found that models with fewer than 20,000 observations in each regression were relatively unstable. To ensure that observations in the model are similar in soil texture, we combine the soil texture groups based off "closeness" of these texture groups. This closeness was determined by the distance between the mean values of silt and clay percentages that define the texture groups. For instance, the "clayey" five group texture classification has a mean proportion of 22.3% silt and 58.6% clay and the medium class has a mean proportion of 25.5% silt and 13.6% clay. This means that the two groups have a distance of $(0.223 - 0.255)^2 + (0.586 - 0.136)^2 = 0.204$. Table A4 shows the soil texture distances. If a soil group had less than 20,000 observations within an MLRA, it was iteratively combined with the next closest group until this observational threshold was reached.

Table A3: Soil Texture Classifications

Group 12 Desig.	Group 5 Desig.
Sand	Sandy
Loamy Sand	Sandy
Sandy Loam	Moderately Sandy
Silt	Medium
Silt Loam	Medium
Loam	Medium
Clay Loam	Moderately Clayey
Sandy Clay Loam	Moderately Clayey
Silty Clay Loam	Moderately Clayey
Sandy Clay	Clayey
Silty Clay	Clayey
Clay	Clayey

Table A4: Soil Texture Group Distances

Texture Group	Clayey	Medium	Mod. Clayey	Mod. Sandy	Sandy
Clayey	0	0.204	0.104	0.416	0.681
Medium	–	0	0.0445	0.158	0.365
Mod. Clayey	–	–	0	0.113	0.286
Mod. Sandy	–	–	–	0	0.0435
Sandy	–	–	–	–	0

Land Use Classifications

Table A5: Cropland Data Layer Observation Designations

Corn	Sweet Potatoes	Greens	Squash
Cotton	Triticale	Herbs	Strawberries
Rice	Alfalfa	Honeydew Melons	Sugarcane
Soybeans	Almonds	Lettuce	Sweet Corn
Spring Wheat	Apples	Mint	Switchgrass
Winter Wheat	Apricots	Misc Vegs	Tobacco
Barley	Aquaculture	Nectarines	Tomatoes
Buckwheat	Asparagus	Olives	Turnips
Camelina	Blueberries	Onions	Vetch
Canola	Broccoli	Oranges	Walnuts
Dry Beans	Cabbage	Other Crops	Watermelons
Durum Wheat	Caneberries	Other Hay/Non Alfalfa	Barren
Fallow/Idle Cropland	Cantaloupes	Other Tree Crops	Clouds/No Data
Flaxseed	Carrots	Peaches	Deciduous Forest
Hops	Cauliflower	Peanuts	Developed (All Levels)
Lentils	Celery	Pears	Evergreen Forest
Millet	Cherries	Peas	Forest
Mustard	Chick Peas	Pecans	Grassland/Pasture
Oats	Christmas Trees	Peppers	Herbaceous Wetlands
Other Small Grains	Citrus	Pistachios	Mixed Forest
Potatoes	Clover/Wildflowers	Plums	Nonag/Undefined
Rape Seed	Cranberries	Pomegranates	Open Water
Rye	Cucumbers	Pop or Orn Corn	Perennial Ice/Snow
Safflower	Eggplants	Prunes	Shrubland
Sorghum	Fruits	Pumpkins	Water
Speltz	Garlic	Radishes	Wetlands
Sugar Beets	Gourds	Shrubland	Woody Wetlands
Sunflower	Grapes	Sod/Grass Seed	
Legend			
Priced Corn Alt.	Corn Alt.	Remaining Crops	Non-Cropland

Logit Coefficient Summary Tables

Table A6: Other-Corn Markov Transition Regression Coefficient Summary

Coefficient	Min	1st Quartile	Median	Mean	3rd Quartile	Max
Intercept	-215.24***	-105.129***	-57.848***	-62.193***	1.854	146.11***
	(71.35)	(12.58)	(13.549)	(13.943)	(15.54)	(29.57)
Corn Price	-1.948***	0.069*	0.312***	0.266***	0.532***	1.615***
	(0.145)	(0.034)	(0.044)	(0.044)	(0.06)	(0.19)
Other Price	-4.889***	-1.165***	-0.643***	-0.476***	-0.044	3.944***
	(0.812)	(0.205)	(0.154)	(0.153)	(0.098)	(0.547)
Slope	-0.544***	-0.055***	-0.016***	-0.001	0.044***	0.196***
	(0.05)	(0.002)	(0.002)	(0.002)	(0.002)	(0.063)
Precip. Q1	-0.791***	-0.119**	0.016	0.031	0.158***	1.097***
	(0.126)	(0.026)	(0.021)	(0.02)	(0.023)	(0.243)
Precip. Q3	-1.221***	-0.199***	-0.104***	-0.089***	0.026	0.732***
	(0.158)	(0.017)	(0.017)	(0.014)	(0.02)	(0.077)
NCCPI Soil Index	-9.948***	0.362***	1.067***	1.197***	1.748***	4.956***
	(0.241)	(0.083)	(0.033)	(0.031)	(0.05)	(0.256)
Irrigation Status	-1.121***	-0.068	0.294***	0.218***	0.542***	2.642***
	(0.958)	(0.124)	(0.114)	(0.106)	(0.09)	(0.09)

Note: Standard errors shown in parentheses. Asterisks *** ,** , and * indicate significance level at 1%, 5%, and 10% respectfully.

Table A7: Corn-Corn Markov Transition Regression Coefficient Summary

Coefficient	Min	1st Quartile	Median	Mean	3rd Quartile	Max
Intercept	-288.347***	7.557	52.25***	51.113***	99.114***	253.779***
	(71.35)	(12.58)	(13.549)	(13.943)	(15.54)	(29.57)
Corn Price	-0.25*	0.134***	0.302***	0.314***	0.46***	1.056***
	(0.145)	(0.034)	(0.044)	(0.044)	(0.06)	(0.19)
Other Price	-3.995***	-1.131***	-0.66***	-0.634***	-0.023	1.435***
	(0.812)	(0.205)	(0.154)	(0.153)	(0.098)	(0.547)
Slope	-0.416***	-0.037***	-0.005**	0.001	0.024***	0.23***
	(0.05)	(0.002)	(0.002)	(0.002)	(0.002)	(0.063)
Precip. Q1	-0.572***	-0.057**	0.024	0.02	0.096***	0.565**
	(0.126)	(0.026)	(0.021)	(0.02)	(0.023)	(0.243)
Precip. Q3	-0.515***	-0.096***	-0.023	-0.018	0.065***	0.457***
	(0.158)	(0.017)	(0.017)	(0.014)	(0.02)	(0.077)
NCCPI Soil Index	-4.783***	-0.826***	-0.402***	-0.421***	0.242***	3.042***
	(0.241)	(0.083)	(0.033)	(0.031)	(0.05)	(0.256)
Irrigation Status	-2.531***	-0.198	0.046	-0.044	0.277***	1.423***
	(0.958)	(0.124)	(0.114)	(0.106)	(0.09)	(0.09)

Note: Standard errors shown in parentheses. Asterisks *** ,** , and * indicate significance level at 1%, 5%, and 10% respectfully.

State Probability Elasticity Summary

Table A8: Summary of State Probability Elasticities

Statistic	Min	1st Quartile	Median	Mean	3rd Quartile	Max
State Prob.	0.097***	0.329***	0.462***	0.432***	0.56***	0.738***
	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)
SR Corn Elast	−1.516***	0.226***	0.476***	0.673***	0.877***	4.857***
	(0.067)	(0.066)	(0.074)	(0.074)	(0.081)	(0.93)
LR Corn Elast	−2.076***	0.067	0.205***	0.51***	0.616***	4.534***
	(0.047)	(0.046)	(0.065)	(0.065)	(0.052)	(0.738)
SR Other Elast	−4.52***	−0.646***	−0.367***	−0.568***	−0.022	1.501***
	(0.078)	(0.073)	(0.087)	(0.087)	(0.083)	(0.258)
LR Other Elast	−7.074***	−0.475***	−0.127**	−0.482***	0	1.582***
	(0.061)	(0.062)	(0.086)	(0.086)	(0.068)	(0.366)

Note: Standard errors shown in parentheses. Asterisks *** ,** , and * indicate significance level at 1%, 5%, and 10% respectfully.

Rotational Marginal Effects, Probabilities, and Elasticities

Table A9: Rotational Estimated Probabilities and Marginal Effects

Statistic	Min	Q1	Median	Mean	Q3	Max
$\mathbb{P}^{CC}$	0.014***	0.053***	0.107***	0.143***	0.202***	0.517***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.005)
$\mathbb{P}^{OO}$	0.022***	0.091***	0.187***	0.278***	0.43***	0.823***
	(0.001)	(0.003)	(0.003)	(0.003)	(0.006)	(0.006)
$\mathbb{P}^{OC}$	0.067***	0.221***	0.312***	0.289***	0.372***	0.426***
	(0.003)	(0.003)	(0.003)	(0.001)	(0.002)	(0.003)
$\frac{\partial \mathbb{P}^{CC}}{\partial P^C}\|_{LR}$	−0.137***	0.008***	0.028***	0.042***	0.075***	0.218***
	(0.049)	(0.002)	(0.003)	(0.005)	(0.009)	(0.035)
$\frac{\partial \mathbb{P}^{CC}}{\partial P^C}\|_{SR}$	−0.032	0.008***	0.017***	0.024***	0.041***	0.121***
	(0.022)	(0.002)	(0.003)	(0.004)	(0.006)	(0.009)
$\frac{\partial \mathbb{P}^{OO}}{\partial P^C}\|_{LR}$	−0.386***	−0.083***	−0.032**	−0.052***	−0.01	0.289*
	(0.135)	(0.015)	(0.013)	(0.012)	(0.007)	(0.148)
$\frac{\partial \mathbb{P}^{OO}}{\partial P^C}\|_{SR}$	−0.225***	−0.049***	−0.022***	−0.03***	−0.007**	0.152***
	(0.04)	(0.005)	(0.004)	(0.004)	(0.003)	(0.034)
$\frac{\partial \mathbb{P}^{OC}}{\partial P^C}\|_{LR}$	−0.076**	−0.015**	0	0.005	0.018***	0.154***
	(0.037)	(0.006)	(0.005)	(0.004)	(0.006)	(0.045)
$\frac{\partial \mathbb{P}^{OC}}{\partial P^C}\|_{SR}$	−0.074***	−0.011***	0.001	0.003	0.013***	0.094***
	(0.019)	(0.003)	(0.002)	(0.002)	(0.003)	(0.017)
$\frac{\partial \mathbb{P}^{CC}}{\partial P^O}\|_{LR}$	−0.485***	−0.165***	−0.059***	−0.086***	−0.002	0.296***
	(0.079)	(0.027)	(0.016)	(0.017)	(0.009)	(0.096)
$\frac{\partial \mathbb{P}^{CC}}{\partial P^O}\|_{SR}$	−0.276***	−0.085***	−0.036***	−0.049***	−0.002	0.181**
	(0.02)	(0.017)	(0.01)	(0.011)	(0.006)	(0.087)
$\frac{\partial \mathbb{P}^{OO}}{\partial P^O}\|_{LR}$	−0.359	−0.046	−0.004	−0.013	0.046	0.203
	(0.225)	(0.031)	(0.035)	(0.036)	(0.04)	(0.462)
$\frac{\partial \mathbb{P}^{OO}}{\partial P^O}\|_{SR}$	−0.221***	−0.04***	−0.006	−0.007	0.031**	0.187*
	(0.065)	(0.011)	(0.013)	(0.01)	(0.013)	(0.097)
$\frac{\partial \mathbb{P}^{OC}}{\partial P^O}\|_{LR}$	−0.589***	0.001	0.064***	0.113***	0.186***	0.897***
	(0.111)	(0.015)	(0.016)	(0.012)	(0.016)	(0.064)
$\frac{\partial \mathbb{P}^{OC}}{\partial P^O}\|_{SR}$	−0.403***	0.004	0.04***	0.062***	0.117***	0.521***
	(0.043)	(0.006)	(0.007)	(0.005)	(0.008)	(0.031)

Note: Standard errors shown in parentheses. Asterisks *** ,** , and * indicate significance level at 1%, 5%, and 10% respectfully.

Table A10: Rotational Elasticities

Statistic	Min	Q1	Median	Mean	Q3	Max
$\frac{\partial \Pi^{CC}}{\partial P^{C}} \|_{LR}$ Elas.	−2.106***	0.633***	1.389***	1.644***	2.32***	7.378***
	(0.672)	(0.133)	(0.15)	(0.198)	(0.293)	(1.322)
$\frac{\partial \Pi^{CC}}{\partial P^{C}} \|_{SR}$ Elas.	−1.021**	0.39***	0.871***	0.953***	1.506***	3.506***
	(0.505)	(0.1)	(0.121)	(0.142)	(0.192)	(0.48)
$\frac{\partial \Pi^{OO}}{\partial P^{C}} \|_{LR}$ Elas.	−4.718**	−2.347***	−1.152***	−1.315***	−0.383	3.575**
	(1.977)	(0.421)	(0.312)	(0.238)	(0.233)	(1.538)
$\frac{\partial \Pi^{OO}}{\partial P^{C}} \|_{SR}$ Elas.	−3.638***	−1.355***	−0.621***	−0.795***	−0.163**	2.861***
	(0.762)	(0.159)	(0.121)	(0.086)	(0.068)	(0.738)
$\frac{\partial \Pi^{OC}}{\partial P^{C}} \|_{LR}$ Elas.	−2.016*	−0.286***	0.001	0.179***	0.298***	4.249***
	(1.197)	(0.108)	(0.084)	(0.065)	(0.091)	(1.212)
$\frac{\partial \Pi^{OC}}{\partial P^{C}} \|_{SR}$ Elas.	−1.963***	−0.215***	0.013	0.092***	0.21***	2.372***
	(0.59)	(0.05)	(0.039)	(0.027)	(0.039)	(0.441)
$\frac{\partial \Pi^{CC}}{\partial P^{O}} \|_{LR}$ Elas.	−8.311***	−1.923***	−0.909***	−1.314***	−0.068	2.326***
	(1.833)	(0.344)	(0.221)	(0.261)	(0.174)	(0.846)
$\frac{\partial \Pi^{CC}}{\partial P^{O}} \|_{SR}$ Elas.	−4.674***	−1.23***	−0.64***	−0.72***	−0.024	1.925***
	(0.94)	(0.239)	(0.141)	(0.174)	(0.101)	(0.511)
$\frac{\partial \Pi^{OO}}{\partial P^{O}} \|_{LR}$ Elas.	−4.674***	0.03	0.802***	0.891***	1.653***	4.86**
	(1.727)	(0.265)	(0.303)	(0.215)	(0.318)	(2.315)
$\frac{\partial \Pi^{OO}}{\partial P^{O}} \|_{SR}$ Elas.	−3.805***	0.031	0.476***	0.517***	1.074***	2.752***
	(0.818)	(0.086)	(0.11)	(0.073)	(0.124)	(0.632)
$\frac{\partial \Pi^{OC}}{\partial P^{O}} \|_{LR}$ Elas.	−5.108**	−0.263***	−0.027	−0.221***	0.237**	1.427*
	(2.003)	(0.084)	(0.089)	(0.085)	(0.104)	(0.836)
$\frac{\partial \Pi^{OC}}{\partial P^{O}} \|_{SR}$ Elas.	−2.464***	−0.251***	−0.028	−0.11***	0.177***	1.398***
	(0.586)	(0.036)	(0.041)	(0.034)	(0.048)	(0.421)

Note: Standard errors shown in parentheses. Asterisks *** ,** , and * indicate significance level at 1%, 5%, and 10% respectfully.