# Physical and Economic Distance in US Soybean Markets

Iván Flores* and Joseph Janzen†

June 15, 2023

**Abstract**

We analyze the shock to US soybean prices in the fall of 2022 caused by drought-induced low water levels on the Mississippi River. We consider how the effects of this shock varied by location relative to observed year-to-year changes in soybean production which were also a function of weather conditions. Both production and supply chain shocks create relative changes in physical and economic distance between soybeans production and end-use locations, especially export terminals at the Gulf of Mexico. We use differences in local soybean prices or basis as a measure of economic distance, we estimate a spatial difference in differences model to understand the relationship between physical and economic distance specifically related to the supply chain shock. Our results show evidence that, on average, physical proximity to the Mississippi river weakened basis by 2.13 cents per bushel of soybeans.

**Keywords:** Basis, Drought, Commodity markets, Soybeans, Spatial price analysis.

*Corresponding Author: Ph.D. Student, Department of Agricultural and Consumer Economics, University of Illinois Urbana-Champaign, 1301 W Gregory Dr., Urbana, IL, 61801, USA. ivanaf2@illinois.edu

†Assistant Professor, Department of Agricultural and Consumer Economics, University of Illinois Urbana-Champaign. jjanzen@illinois.edu

# 1 Introduction

The Mississippi River is a critical trade route in the US Midwest. The Mississippi and its major tributaries connect the major agricultural production area of the United States to ocean transport to major export markets. In October 2022, drought caused Mississippi River water levels to reach their lowest point in a decade. Low water levels disrupted the typical price relationship across space at producing and exporting locations, relative to historical rates particularly in October.

The objective of this paper is to assess the heterogeneous impact of the Mississippi River drought on local soybean markets. Soybeans are a significant part of the US agricultural economy, particularly in terms of exports. As a versatile crop with various uses, including animal feed, cooking oil, and renewable fuel production, shocks to soybean prices touch many downstream sectors both within and outside the US. Soybeans are uniquely reliant on Mississippi River transport relative to other commodities, highlighting the significance of studying the impact of supply chain disruptions on soybean prices. This research seeks to address two key questions; first, how did local soybean prices adapt to the drought-induced changes in production, consumption, and trade flows in the fall of 2022? Second, how did the market response to these shocks compare to the physical proximity of soybean-producing counties to demand sources such as exports that rely on river transportation?

Transaction costs are the key factor that determines price differences across locations for a homogeneous commodity like soybeans. von Cramon-Taubadel and Goodwin (2021) defines the Law of One Price as the rule that price differences between markets should not exceed the transaction costs associated with buying in the relatively cheap market and selling in the relatively expensive one. Thus, differences in soybean basis must reflect differences in transaction costs involved in such arbitrage. When one of these transaction costs changes, like the cost of barge to export points in fall 2022, the actual basis change may reflect not only the transport cost change, but also the relative pull of other use markets such as the domestic crushing market. Typically, these transactions costs are strongly correlated with the physical distance between origin and destination markets.

When a critical route on the commodity supply chain such as the Mississippi River is disrupted, the resulting bottleneck creates differential changes in transactions costs across time and space (Garfinkel and Rao, 1971). In 2022, the significant changes in river water levels affected the width and depth of the river flow, leading to limitations in barge loading and tow-boat carrying capacity and thus the economic costs of moving commodities from supply to demand locations. The width of the river was affected, restricting the number of barges that could move during periods of low water levels. Similarly, the depth of the river prevented barges from being loaded to normal weight thresholds. Figure 1 presents a satellite image of the river near Eudora, Arkansas, where sections with reduced width can be observed, resulting in slower commercial traffic and loading restrictions for Midwest grain and other commodities. Figure 2 displays backed-up barges around Vicksburg, Mississippi, where low water levels forced transit to come to a halt in certain river sections. These dynamics not only impacted transport time but also the maximum loading capacity for grain. The longer time required to transport goods, specifically soybeans on a barge in our context, directly translates into higher transaction costs.

While barge rates increased dramatically to account for higher transactions costs, the true impact of the supply chain disruption of 2022 must also account for behavioral changes by producers, shippers, and end-users in response. Farmers and merchants may opt for alternative means of transportation, explore different destinations, or simply wait and store their grain until more favorable conditions arise. All of these factors introduce non-linearities in the relationship between prices among physical locations of producers and export points, making it more challenging to observe the relevant transaction costs. In these cases, physical distance can no longer serve as a reliable proxy for economic distance. Consequently, our primary focus lies in examining the shift of soybean prices. This perspective enables us to explore the connection between changes in economic distance, influenced by the drought's effects and proximity to river or domestic markets, and the resulting price patterns. It is important to note that the impact of the drought on prices varies across locations due to heterogeneous weather effects. Regions heavily reliant on the river face higher exposure to weather-related shocks, which limits their flexibility and adaptability in dealing with such challenges.

We consider the cash-futures basis across locations as a measure of economic distance. Basis is the difference between local cash price and the nearby futures price and a key piece of information used by soybean producers, traders, and end users when making marketing decisions(Jiang, 1997). For instance, producers compare basis bids from elevators, crushers, or export points when deciding to sell. The closest market to the producer in terms of offering the highest basis may not be the location in nearest physical proximity, even though the transport costs to the nearest location may be lowest.

Previous studies have examined spatial price relationships in similar circumstances. McNew (1996) studied the effects of the Midwest flood of 1993, along the Mississippi River. They find that higher transportation costs caused by the river flood, reduced corn market integration significantly, consequently decreasing the transfer of excess demand shock across regions. The paper also highlighted the significance of non-linearity in spatial price relationships, emphasizing that integration between supply chain points enables perfect transmission of price shocks, whereas lack of integration hampers spatial transmission. Tomek and Kaiser (2014) discuss that in spatial price relationships, the relevant costs are the full costs of spatial arbitrage. In line with von Cramon-Taubadel and Goodwin (2021), they describe that in a competitive market structure, spatial price relationships are determined by transfer costs among regions. Tomek and Kaiser (2014) suggests that inter-regional trading patterns and price relationships can be analyzed under a market integration framework to unveil spatial price variations.

The paper makes three key contributions to our understanding of spatial price analysis. First, the paper uncovers compelling evidence of risk exposure to weather-related shocks that vary depending on spatial location within the soybean supply chain. By examining the spatial heterogeneity of these risks, the research contributes to a deeper understanding of the vulnerability of different regions along the supply chain and provides valuable insights for policymakers and industry stakeholders in devising effective risk management strategies. Secondly, the paper utilizes a spatial difference-in-differences methodology to estimate the impact of river proximity on soybean basis change. By accounting for spatial variations and employing a quasi-experimental design, the study provides robust evidence on the causal relationship between river proximity and changes in soybean basis, shedding light on the spatial dynamics influencing the agricultural market. Lastly, it employs spatial interpolation techniques to generate visualizations that depict a smooth pattern of prices across space, allowing researchers to gain a comprehensive understanding of the shift in patterns from 2021 to 2022. This approach enables the identification of nuanced spatial variations in price changes, aiding in the identification of potential underlying factors driving these shifts.

This paper reviews the spatial distribution of soybean production and the typical movement from production to consumption locations. To analyze the change in soybean prices pattern in this context, we combine unique data on prices, origins and end users destinations. Section 2 reviews relevant information about the organization of US soybean production and consumption. Section 3 describes the price and quantity data used for our analysis, including the geospatial strategies used to associate production and prices. We describe observed differences in physical and economic distance and how both can change over time. In section 5, we outline our empirical procedure for identifying the portion of observed changes in economic distance which are attributable to river transport disruptions. Our approach is analogous to a spatial differences-in-differences research design that consider basis changes across space relative to past basis patterns. We show the impact of low water levels had heterogeneous impacts over space and discuss and the implications of this result for future marketing and policy decisions. We also acknowledge the limitations presented and propose further strategies to our paper.

## 2 Soybean production in the US

In the US, soybeans are mainly grown in the Midwest of the country; the top five producing states are Illinois, Iowa, Minnesota, Nebraska, and Indiana. Figure 3 shows the county-level production data for 2021 (panel A) and 2022 (panel B). Both panels also show the location of processing plants and the three major waterways used for barge transportation, the Mississippi, Ohio, and Illinois Rivers. Much of the US soybean production is in areas connected to river shipping. The location of soybean production is mainly related to conducive soil and climatic conditions found in areas where soybeans are currently produced. Technological developments like the creation of seed varieties adapted to shorter growing seasons have allowed soybean production to expand north and westward into North and South Dakota. Finally, US soybean production possess the the competitive advantage offered by river transport to efficiently move grain to export locations. Counties near the Illinois River display high production as well as locations around the Mississippi River in Tennessee, southeast Missouri, Arkansas and Mississippi. Significant quantities of soybeans are also produced to the west from North Dakota south to Kansas.

Comparing panels in Figure 3 demonstrates the impact of weather on production. Between 2021 and 2022, the largest production changes occurred in western states with significant yield declines in Kansas, Nebraska, and western Iowa.US soybean production in 2021 was 4.46 billion bushels, a record high largely due to favorable weather conditions and high yields. Production in 2022 was 4.28 billion bushels, down 4% from 2021, even though US soybean planted acreage was up slightly from the year prior. These year-to-year fluctuations point to the importance of weather as a supply and demand driver in soybean and other agricultural markets.

Following the harvest, soybeans are directed towards two main destinations: local processors and export terminals. Roughly half of US soybeans are processed domestically and half are exported prior to being processed in other countries. Processing occurs at specialized facilities known as crush plants, which play a vital role in the soybean supply chain. At these plants, soybeans undergo a mechanical extraction process, commonly referred to as "crushing," resulting in the production of two valuable commodities: soybean meal and soybean oil. Soybean meal serves as a crucial ingredient in animal feed, providing essential proteins and nutrients. Soybean oil is used as a food ingredient in cooking and baking, as well as for the production of biodiesel, a renewable transportation fuel.

The placement of crush plants is influenced by several factors including the proximity to soybean production regions and transportation infrastructure. The location of these facilities is crucial to minimize transportation costs and ensure a steady supply of soybeans. Historically, soybean prices have exhibited fluctuations due to various factors such as weather conditions, global demand, and market dynamics. These price fluctuations have a direct impact on the profitability of crush plants. Consequently, the geographical distribution of crush plants often aligns with areas where soybean production is prominent, allowing for cost-effective sourcing of raw materials. Additionally, the proximity to export points facilitates efficient transportation of processed soybean products to global markets, enhancing competitiveness and market access.

The alternative destination of soybeans, export points, are crucial in the soybeans supply chain. We focus on export terminals on the Gulf of Mexico since they are directly connected to the Mississippi river waterway and handle the largest share of US exports. Gulf terminals moved 59% of soybean exports in 2022, compared to 25% for points in the Pacific Northwest, 10% for interior movement via land to Canada and Mexico, 5% for Atlantic and 1% for the Great Lakes export terminals. The primary destinations for soybean exports are China, the European Union and Mexico.

Transportation of soybeans in the US typically involves some combination of truck, rail, and barge transportation depending on origin, destination, and infrastructure availability. Trucks typically transport soybeans from farms to local storage facilities or processing plants. From there, soybeans may be transported by rail to other parts of the country or by barge along major waterways such as the Mississippi River to coastal ports for export. Transportation mode is a function of both distance and destination. On a per-bushel and per-distance basis, truck is typically the most expensive mode of transport, followed by rail and barge. For this reason, truck transport is typically used only for shorter journeys of approximately 100 miles or less.

Given transportation costs and distances, around 60% of soybean production is transported by barge with an average cost of $0.03 to $0.06 per bushel per mile, typically moved into The Gulf of Mexico at harvest. US railroads move around 30% of total production, mostly to inland markets, PNW and The Lakes. The rail average cost of transport is around $0.04 to $0.07 per bushel per mile. Truck move the remaining 10% to closer hauls, with a higher cost that rose $0.20 to $0.40 per bushel per mile.

## 3 Data

This paper uses multiple data sources to describe soybean supply chains. The origin for this supply chain is the producer, so we use county-level soybean production data for 2021 and 2022 from the USDA National Agricultural Statistics Service. The end-point of the supply chain are the locations of domestic crush plants and export terminals. We collect these points from Bloomberg and the National Oil-seed Processors Association. Important intermediate points on the supply chain include barge-loading elevators located along major navigable rivers, specifically the Mississippi, Illinois, and Ohio. To measure the value of soybeans across these locations, we collect spot price data from Bloomberg for local markets, crush plants, and export locations along with benchmark futures price data. We calculate the basis as the difference between the spot price at a given location and the nearby futures price.

Using shapefiles describing the spatial coordinates of US counties, end-use locations, and major rivers, we create a data set with geospatial components that allows us to calculate spatial measures such as physical distance between each centroid of the counties producing soybeans and the points that demand soybeans, crushing plants. Moreover, we are able to interpolate soybean basis across space using an inverse weighted distance process to fill in basis data in space acknowledging the difference generated by physical distance among locations. This process assumes that any meaningful deviations from a consistent spatial gradient of prices over space are observed in the data. Our data provide comprehensive coverage over space that validate this assumption as discussed below.

## 3.1 Production

The soybean production data at the county level provided by USDA-NASS offers valuable insights into the regional variations and contributions of soybean cultivation across the United States. This comprehensive data-set encompasses information on soybean acreage, yield, and total production for each county. We use total county-level production from 2021 and 2022. Since production is more diffuse than end-use locations, we describe the supply chain thinking about production locations. In particular, we take county level production data, that it is the main reason to include county in our data as the main unit of analysis. We have data from 1,229 counties that make up 80% of US soybean production.

## 3.2 Physical distance

To describe the connection between the diffuse set of production locations and the concentrated number of domestic crush plants and export terminals, we calculate the physical distance between them. Physical distance refers to the actual spatial separation in miles between different locations involved in the production, processing, and distribution of soybeans. The process of calculating the physical distance involves calculating the euclidean distance between points in our data. We represent each producing county by its centroid and each crush plant by its latitude and longitude.

Due to relative transportation costs, processing capacity is concentrated away from major rivers. The bulk of domestic soybean processing capacity exists in an arc from Southern Minnesota to Ohio, with nearly all major producing states having some processing capacity. As shown in figure 4 areas with high production levels, such as central Illinois, western Iowa and central Indiana, have a high concentration of local soybean demand with most counties being less than 25 miles from the nearest crush plant. However, new production areas such as North Dakota, are in a range of more than 160 miles from at least one source of local demand. US crush capacity expansion related to the boom in renewable diesel production is expected to focus on the western part of the US soybean growing area defined in figure 3. Other major production areas as defined in Figure 3, especially Illinois and states along the lower Mississippi are geared more toward river transport. Processing facilities are largely absent in these areas, especially in counties adjacent to the Illinois and Mississippi Rivers.

To summarize the distance between the typical bushel of US soybeans and its destination, we calculate the production-weighted average distance between soybean-producing counties and the nearest domestic crush plant and the main export location at the Gulf of Mexico. This quantity is a benchmark for the proximity to market of each bushel of soybeans. For each county, the distance to soybean processing facilities is the straight-line distance from the county centroid to the plant location. The distance to export terminals on the Gulf of Mexico is the straight-line distance from the county to the nearest barge-loading facility on a major river, plus the number of river miles between that facility and the Gulf. Although this is an imperfect measure because it treats overland miles the same as river miles, it recognizes that the importance of the export market to a given location depends on both the ability to access barge transportation and the proximity of that barge loading location to export terminals.

Table 1: Proximity of US Soybeans to End-Use Locations, 2021 versus 2022.

| **Production-weighted average distance (in miles) to:** | **2021** | **2022** | **Change (%)** |
|---|---|---|---|
| Nearest soybean processing facility | 55.2 | 54.6 | -1.1 |
| Gulf export terminals | 1,052.0 | 1,052.2 | 0.0 |

Table 1 shows the average production-weighted physical distance of a bushel of soybeans in counties with NASS production data for both 2021 and 2022. Because production moves from year to year due to changes in acreage and yield over time and across space, the distribution of physical distance to processing facilities and export terminals may also change even though crush plant and export terminal locations do not change much if at all from year to year. For example, if production declines in areas far away from processing capacity, the average physical distance of a bushel of soybeans to the nearest processing plant may fall. Despite observed changes in production and yield, the average distance of a bushel of US soybeans to crush plants or export terminals changed little between 2021 and 2022. The average distance to a crush plant fell from 55.1 miles to 54.6 miles, or just 1.1%. Similarly, observed production declines in Iowa, Kansas, and Nebraska, locations relatively far from river transportation and thus export terminals, did not reduce the average distance of a bushel of US soybeans to Gulf export locations, which rose slightly from 1,052.0 miles to 1052.2 miles. It appears that production decreases in these far-from-market locations were offset by increases in other distant production locations, especially North Dakota.

## 3.3 Prices

When it comes to the spatial distribution of prices, the pattern tends to show higher prices (and the spot-Gulf difference is smaller) at locations closer to the Gulf. Figure 5 shows the spatial distributions of local spot prices relative to Gulf export bids in 2021 and 2022. Locations near major rivers with cheaper access to barge transportation also have relatively smaller price discounts relative to the Gulf. Figure 5, panel A shows the spatial distribution of October 2021 price differences where values are higher for locations further south and closer to major rivers. Pockets of higher prices can also be seen near crush plants locations. The distribution of prices in October 2022 shows how high transportation costs led to abnormally large price differences between inland locations and the Gulf, especially along major rivers. Price differences in 2022 are larger than in 2021 and concentrated along the lower Mississippi and Ohio Rivers. Markets adjusted to discourage soybean movement to the river system, temporarily favoring domestic processing.

In Figure 6, we show a visual representation of how local markets reacted to conditions in the fall of 2022. This visualization focuses on the change in the price difference between the spot price and the Gulf price between 2021 and 2022, which can be observed by comparing Panels A and B in Figure 6. The most significant variations are evident in regions near major rivers, specifically along the lower Mississippi River and in southern parts of Illinois, Indiana, Kentucky, and Missouri. In 2022, the spot-Gulf price difference was \$2.20 per bushel lower compared to 2021. Conversely, the areas that experienced the greatest impact from drought, such as Kansas and Nebraska, exhibited the smallest changes. In these regions, the spot-Gulf price difference was around \$0.60 to \$1.00 per bushel lower.

The basis interpolation was based on the inverse weighted distance procedure. Given a set of observed data points $(x_i, y_i, z_i)$ with known values $z_i$ at coordinates $(x_i, y_i)$ for $i = 1$ to $n$ is considered. To estimate the value at a target location $(x_0, y_0)$, the inverse weighted distance, denoted as $w_i$, is calculated for each observed point. This weight is determined by taking the reciprocal of the distance between the observed point and the target location raised to a power p: $w_1 = 1/(distance((x_0, y_0), (x_i, y_i)))^p$. Here, the distance function measures the Euclidean distance between the target location and the observed point. The interpolated value at the target location, denoted as $z_0$, is obtained through a weighted averaging process of the known values at the observed points: $z_0 = \sum(w_i * z_i)/\sum w_i$. The summation is performed over all observed points, where the numerator calculates the weighted sum of the known values, and the denominator computes the sum of the weights. Adjusting the power parameter p allows for the manipulation of neighboring point influence. Higher values of p assign more weight to nearby points, resulting in a smoother interpolation, while lower values enable more distant points to contribute, leading to a more variable interpolated surface.

### 3.4 Economic distance

Transaction costs are hard to observe given the changes produced to supply, demand, and carrying costs generated by the drought shock. Physical distance is an imperfect proxy for economic distance. When it comes to prices however, as pointed by (von Cramon-Taubadel and Goodwin, 2021), price differences among markets at different locations measure the economic distance between those points. The law of one price underscores the economic function of local spot markets to encourage efficient commodity movement from places where the commodity is abundant to where it is scarce. Large price differences between locations suggest that such movement is expensive, and the two locations are far from each other in economic terms, even if they are relatively close in terms of physical distance.

We consider two measures of relative prices among soybean markets in the US: the spot-futures basis and the spot-Gulf price difference. The basis describes the spot price, the value of a bushel of soybeans at a given location, relative to the benchmark price level given by the nearby futures price. The inland spot-Gulf price difference is the discount in each local market relative to the price at export terminals on the Gulf of Mexico. We collect data on futures prices, export bids at the Gulf, and local spot prices for 2,944 elevators, crush plants, and other locations throughout major US soybean growing areas from Bloomberg. To eliminate seasonal price patterns related to the returns to commodity storage, we focus on harvest-time prices represented by the average of observed prices during the month of October in 2021 and 2022.

Table 2: Soybean Price Levels and Average Differences Between Spot and Benchmark Prices, October 2021 versus October 2022.

| | **2021** | **2022** | **Change** |
|---|---|---|---|
| **Price levels (in $/bushel)** | | | |
| Nearest futures price | 12.30 | 13.81 | +1.51 |
| Gulf export bid | 13.19 | 15.92 | +2.73 |
| Difference | +0.89 | +2.11 | |
| **Production-weighted average price difference** | | | |
| Spot-futures | -0.44 | -0.53 | -0.09 |
| Spot-Gulf | -1.29 | -2.63 | -1.34 |

Table 2 illustrates the dramatic combined impact of drought, river transport disruptions, and strong demand for US soybeans on prices observed in fall 2022. Soybean price levels were higher in fall 2022 than in fall 2021, but the price change was especially pronounced at Gulf export terminals. The average nearby futures price was $13.81/bushel, or $1.51/bushel higher than the year prior, while Gulf export bids were $15.92/bushel, or $2.73 higher than the year prior. Table 2 also shows that values at inland locations did not increase nearly as much as the price at the Gulf; the economic distance between these points widened to account for higher transportation costs. To summarize this change, we calculate the production-weighted average across all locations of the spot-futures and spot-Gulf price differences. Production-weighting ensures that we do not overweight areas with relatively high or low prices (that is, relatively strong or weak basis) but little production. Average spot-futures basis weakened by 9 cents per bushel, while the average spot-Gulf price difference was $1.34/bushel lower.

The price data collected considers multiple harvest periods; pre-2022 data help to form the baseline against which we can compare what occurred in fall 2022. Since the bid data is at the elevator/crush level, we aggregate the data for each county so that there is one value per unit of observation in the regression.

Our main unit of observation is county-year specific. We use county-level soybean production to control that the change in basis is not mediated by a production shock, but solely by river proximity variation. Production in 2021 and 2022 is relatively constant although, showing only a 4% decrease in total production from one year to the other. The available data from USDA does not consider the same counties for 2021 and 2022, so counties that did not had production data in 2021 but had in 2022 or vice-versa were dropped from our sample.

Table 3 presents a comprehensive summary of statistics for a set of variables in the regression equation. It includes information on counts, means, standard deviations, minimum and maximum values for each variable. The variables encompass various aspects such as soybean basis values for two different years, a binary variable indicating treatment status, soybean production figures for two years, distances to different locations, and changes in basis and production over the given period. The table provides a valuable overview of the statistical characteristics of these variables, aiding in the understanding and analysis of the data at hand.

Table 3: Summary of statistics

| Variable | Count | Mean | St. Dev. | Min | Max |
|---|---|---|---|---|---|
| Soybean basis 2021 | 1,229 | −0.452 | 0.208 | −1.115 | 0.206 |
| Soybean basis 2022 | 1,229 | −0.595 | 0.318 | −1.424 | 0.279 |
| Treated (river adjacent county = 1) | 1,229 | 0.107 | 0.309 | 0 | 1 |
| Soybean production 2021 | 1,229 | 2,664,913 | 3,151,127 | 0 | 20,460,000 |
| Soybean production 2022 | 1,229 | 2,572,993 | 3,100,218 | 0 | 21,186,000 |
| Distance to crusher | 1,229 | 96.699 | 68.333 | 1 | 312 |
| Distance to Gulf (along the river) | 1,229 | 1,002.522 | 290.086 | 369 | 1,682 |
| Distance to river elevator | 1,229 | 162.337 | 132.188 | 0 | 597 |
| Distance to Gulf (Euclidean distance) | 1,229 | 840.185 | 254.447 | 362 | 1,288 |
| $\Delta$ basis | 1,229 | −0.143 | 0.354 | −1.075 | 0.819 |
| $\Delta$ production | 1,229 | −91,920 | 582,914 | −3,438,000 | 3,960,000 |

# 4 Methods: Spatial Difference in Differences

In order to assess the influence of reduced Mississippi River levels on basis, we employ a spatial difference-in-difference methodology. This approach involves comparisons over time among locations that vary in their proximity to domestic crush and export demand, thus experiencing different degrees of exposure to the river shock observed in 2022. By examining data both before and during the 2022 harvest period, which was affected by low water conditions, we can capture the spatial dynamics of the soybean market and identify the extent to which counties were subject to treatment effects during the pre- and post-drought periods. This approach accounts for the common impact of the river drought and other market shocks on basis in all counties as well as the unique attributes of each county that lead to basis patterns that do not vary over time.

Proximity is flexibly defined since it will depend on the different market substitutes a county has access to. In this case we do not impose any distance threshold, but let the coefficients show how distance to a river or to a crusher generates an impact on basis. It is important to notice that changes in observed prices are not representing the river drought effect nor the price difference over space alone.

The Difference-in-Differences (DiD) approach is a widely used method in econometrics to estimate causal effects (Angrist and Pischke, 2009). It involves comparing the changes in outcomes between a treatment and a control groups before and after an intervention. This method relies on the assumption that, in the absence of the treatment, the trends in the outcome variables for the treatment and control groups would have been parallel. By exploiting this parallel trend assumption, the DiD approach aims to isolate the causal effect of the treatment.

The Spatial Difference-in-Differences is an extension of the traditional Difference-in-Differences (DiD) approach that incorporates spatial dimensions into the analysis. In particular, it acknowledges the presence of spatial relationships in the dependent variable, exogenous variables and error term (Kelejian and Prucha, 2010). The spatial DiD is particularly useful when studying interventions or shocks that have spatially varying effects, combining the temporal and spatial dimensions to estimate the causal impact of a treatment on an outcome variable.

One of the key advantages of the spatial approach is its ability to account for spatial spillover effects. Even when we do not explicitly measure the spillover effects, we recognize that the treatment effect may extend beyond the treated area and affect neighboring regions. By explicitly considering the spatial dimension, it allows for the identification of both direct and indirect effects of the treatment. Additionally, the DiD approach can handle situations where the treatment assignment is based on non-random selection.

Our underlying identifying assumption is that in absence of drought, basis would have changed similarly between 2021 and 2022 at locations close to the river and the gulf, as it did at far away locations, conditional on production shocks and proximity to crush plants.

To test the change in basis we estimated the following regression:

$$\Delta \text{ Basis}_i = \gamma \, \text{Exposure}_i + \theta \, \Delta \text{ Production}_i + \phi \, \text{Crush Distance}_i + \varepsilon_i \tag{1}$$

Where Δ Basis is the change in basis in county $i$ from 2021 to 2022. Exposure is measure of exposition of county's $i$ to the river. It considers two measures of exposure, which are land distance from county $i$ centroid to its nearest river elevator, and distance along the river from the elevator to The Gulf. Δ Production is the change in soybean production in county $i$ from 2021 to 2022. Crush Distance is the euclidean distance from county $i$ centroid to its nearest Crusher.

## 5 Results and discussion

Table 4 presents coefficient estimates for equation 1 where the dependent variable is the change in basis (Δbasis) between 2021 and 2022. The analysis considers several independent variables related to distance[1] and production.

The results indicate that distance to the river elevator has a significant positive effect on the change in basis. In Model 1, for every hundred miles away from a river elevator, the change in basis increases by 0.113 units ($p < 0.01$). This positive relationship holds in Model 2 (0.160, $p < 0.01$) and Model 3 (0.375, $p < 0.01$) as well.

Similarly, distance to the Gulf (along the river) also has a positive effect on the change in basis. In Model 1, a hundred miles increase in distance away from the Gulf results in a 0.026 unit increase in the change in basis ($p < 0.01$). This effect remains significant in Model 2 (0.013, $p < 0.01$) and Model 3 (0.106, $p < 0.01$).

The squared terms of distance to the river elevator and distance to the Gulf (along the river) show negative effects on the change in basis. In Model 3, the squared term for distance to the river elevator has a significant negative effect (-0.047, $p < 0.01$), indicating a non-linear relationship between distance and basis change.

Additionally, the distance to the crusher shows a significant negative effect on the change in basis in Model 2 (-0.246, $p < 0.01$) and Model 3 (-0.213, $p < 0.01$), suggesting that every hundred miles away from a crusher is associated with a decrease in basis.

The change in production does not show a significant effect on the change in basis.

[1]Note: River miles are calculated through the course of the Mississippi River, while land miles are calculated as the euclidean distance between the centroid of a county and the nearest river elevator. Although, the measure of miles is not identical it provides a very good approximation to the real world supply chain structure

Table 4: Relationships between Physical Distance and Changes in Soybean Basis

| | Dependent variable: $\Delta basis$ | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Distance to river elevator | 0.110*** | 0.148*** | 0.358*** |
| | (0.007) | (0.007) | (0.018) |
| River Distance to Gulf | 0.029*** | 0.018*** | 0.106*** |
| | (0.004) | (0.003) | (0.020) |
| Distance to river elevator$^2$ | | | −0.045*** |
| | | | (0.004) |
| River Distance to Gulf$^2$ | | | −0.006*** |
| | | | (0.001) |
| Distance to Crusher | | −0.219*** | −0.175*** |
| | | (0.013) | (0.039) |
| Distance to Crusher$^2$ | | | −0.006 |
| | | | (0.015) |
| $\Delta$ Production | | −0.000*** | −0.000*** |
| | | (0.000) | (0.000) |
| Observations | 1,229 | 1,229 | 1,229 |
| Adjusted $R^2$ | 0.216 | 0.411 | 0.480 |

*Note: Distance in hundred miles.* *p<0.1; **p<0.05; ***p<0.01

*Robust standard errors in parenthesis.*

The results from model 3 provide valuable insights into the implications of basis changes for soybeans transportation costs. The change in basis, which serves as a measure of how much the economic distance has changed between supply and demand points, is influenced by various factors, including distance to river elevators, distance to the Gulf (along the river), and distance to the crusher.

The coefficient estimate for the distance to river elevator squared term (-0.045) suggests a non-linear relationship between distance and basis change. As the distance to the river elevator increases, the rate of increase in basis change slows down. This implies that basis tend to rise at a decreasing rate as the distance to the river elevator increases.

The coefficient estimate for the distance to the Gulf squared term (-0.006) also indicates a non-linear relationship. Similar to the distance to the river elevator, as the distance to the Gulf (along the river) increases, the rate of increase in basis slows down. Again, this suggests that basis change rise at a decreasing rate as the distance to the Gulf increases.

These findings have important implications for understanding the cost structure of soybean transportation during the fall of 2022 when river drought increased river transportation costs. Given that around 60% of soybean production is transported by barge into the Gulf of Mexico, the results indicate that the costs per bushel per mile for barge transportation are likely to increase at a decreasing rate as the distance from the river elevators or the Gulf of Mexico increases. This implies that the cost per bushel of soybeans for county a hundred miles away from a river elevator would decrease by \$1.07 to \$2.15. This cost would decrease from \$0.32 to \$0.64 for every hundred miles away from The Gulf. A county hundred miles away from a crusher would experience a increase in transportation costs of about \$0.53 to \$1.05 per bushel of soybeans.

The results indicate that transportation costs for soybeans increase with closeness to the river, but at a decreasing rate. This suggests that longer transportation distances dependent on barges would be mostly impacted by the drought. It is important to note that these implications are based on the back-of-the-envelope calculations using the coefficient estimates from Model 3 and the provided information on transportation costs. Further analysis and consideration of other factors are necessary for more precise estimations.

# 6 Conclusion and further work

Our results provide valuable insights into the impact of the Mississippi River drought on soybean basis during the fall of 2022. Through a visual analysis of the basis patterns compared to the previous year (2021), as well as a regression analysis, we examine how the relative distance to the river and the Gulf influenced the change in soybean basis for counties producing soybeans.

The regression analysis reveals several significant findings. Firstly, the distance to the river elevator has a positive effect on the change in basis, indicating that as the distance to the river elevator increases, the basis also increases. This positive relationship holds across all three models.

Similarly, the distance to the Gulf (along the river) has a positive effect on the change in basis. Counties located farther from the Gulf experienced larger increases in basis. This effect remains consistent in all three models.

The squared terms for distance to the river elevator and distance to the Gulf exhibit negative effects on the change in basis. This suggests a non-linear relationship between distance and basis change. Specifically, as the distance to the river elevator or the Gulf increases beyond a certain point, the rate of increase in basis slows down.

Furthermore, the distance to the crusher shows a significant negative effect on the change in basis. Counties far away to a crusher experience a decrease in basis. This finding highlights the impact of the crushers as local market demand in October of 2022, opposed as the traditional pattern for this month of the Soybean marketing calendar.

Notably, the change in production does not significantly affect the change in basis, indicating that other factors, such as distance, have a more prominent role in influencing basis fluctuations.

These findings contribute to our understanding of the relationship between distance, production, and basis changes in the context of the Mississippi River drought.

Further work includes adding to our model a measure of heterogeneity along the river as well as including how much a foot of water level impact soybean basis.

## References


Angrist, J. D. and Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Garfinkel, R. S. and Rao, M. (1971). The bottleneck transportation problem. *Naval Research Logistics Quarterly*, 18(4):465–472.

Jiang, B. (1997). *Corn and soybean basis behavior and forecasting: fundamental and alternative approaches*. Iowa State University.

Kelejian, H. H. and Prucha, I. R. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of econometrics*, 157(1):53–67.

McNew, K. (1996). Spatial market integration: Definition, theory, and evidence. *Agricultural and Resource Economics Review*, 25(1):1–11.

Tomek, W. G. and Kaiser, H. M. (2014). *Agricultural product prices*. Cornell University Press.

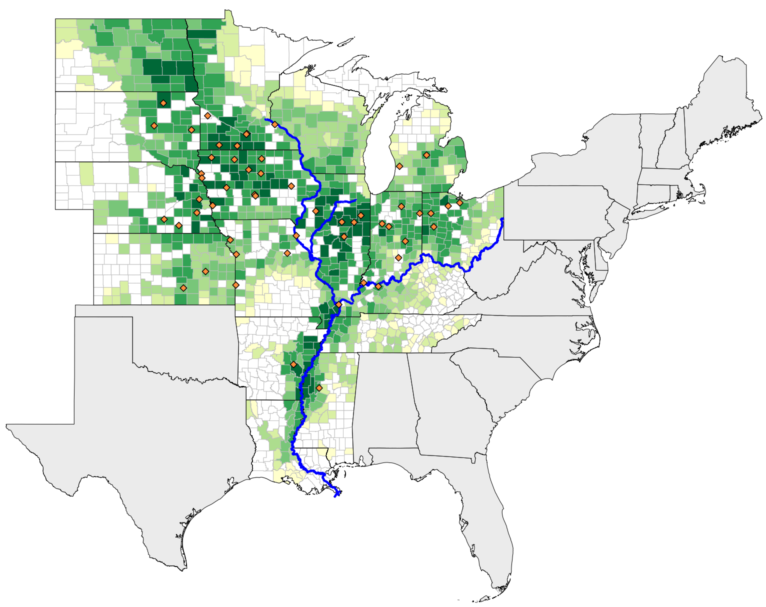
von Cramon-Taubadel, S. and Goodwin, B. K. (2021). Price transmission in agricultural markets. *Annual Review of Resource Economics*, 13:65–84.

Appendix

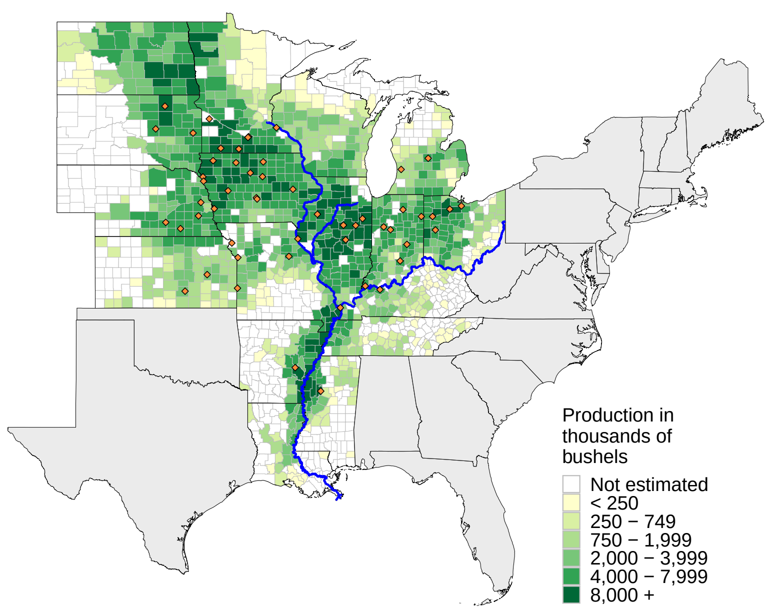
# A Figures


Figure 1: Mississippi River near Eudora, AR. October 2021 and 2022. *Source: European Union Copernicus Sentinel-3 imagery*


Figure 2: Backed-up Barges North of Vicksburg, Mississippi. October 7, 2022. *Source: European Union Copernicus Sentinel-3 imagery 2*



(a) October, 2021




(b) October, 2022

Figure 3: US County-level Production by Year and Crush Plant Location.

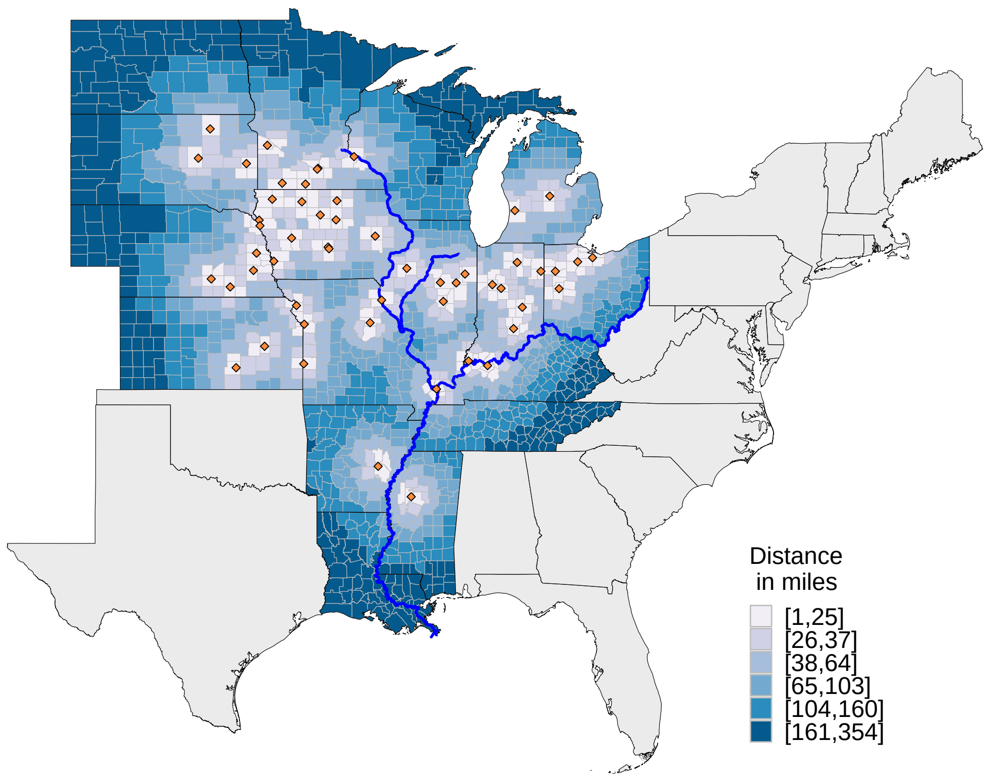
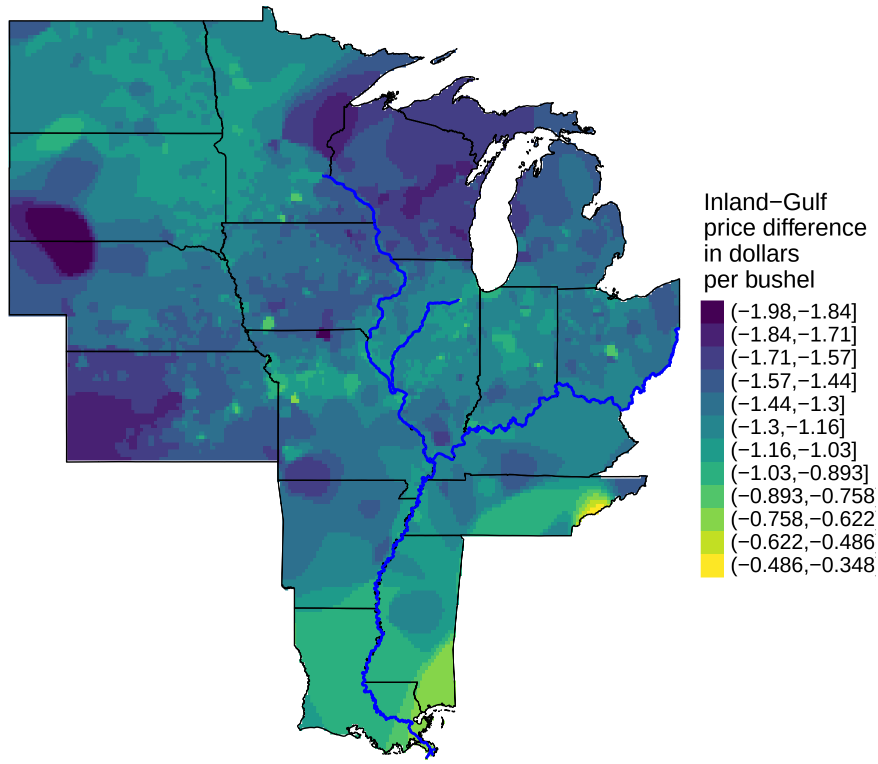
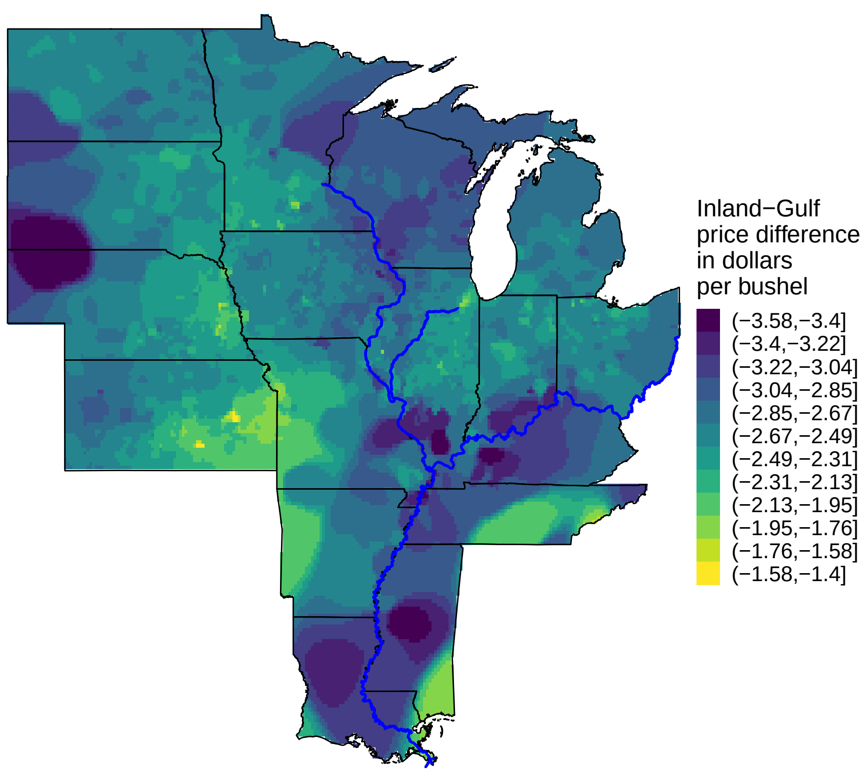



Figure 4: Physical Distance, Minimum Distance from Counties to Local Markets




(a) October, 2021






(b) October, 2022

Figure 5: Differences Between Soybean Spot Prices and Gulf Export Bids by Location.

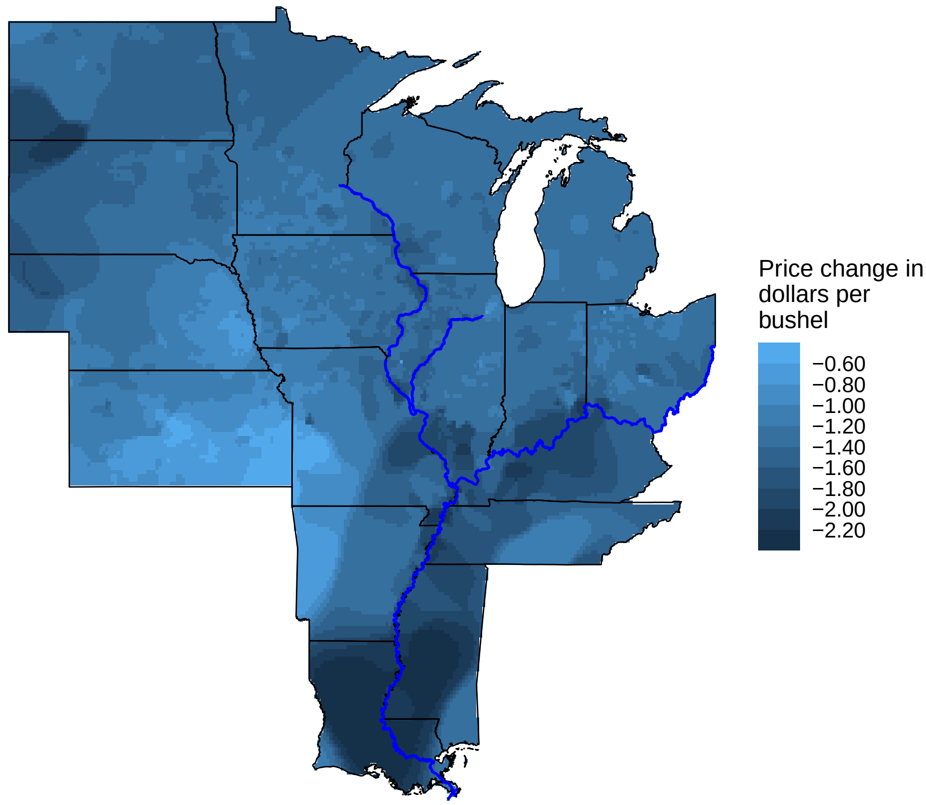



Figure 6: Change in October Spot-Gulf Price Difference by Location, 2022 minus 2021.