

Globally Distributed Production and the Pricing of CME Commodity Futures

June 12, 2013

Abstract

I investigate how local supply shocks in the globally distributed production of commodities are incorporated into CME futures prices. I exploit that the soybean market share of the US (Argentina) decreased (increased) between 1996 and 2010, and use rain, which tends to increase output, as a source of exogenous supply shocks. I find a significantly negative response of CME soybean prices to daily rain across regions and time. Moreover, the impact of local rain in the CME price is approximately linear in the time-varying local share of global output. Therefore, traders of CME contracts seem to aggregate supply in a globally integrated manner and are exposed to globally distributed shocks.

Keywords: Commodities; Futures; Weather; Supply Shocks

JEL codes: G13; G14; Q11.

1. Introduction

In this paper I investigate how geographically distributed supply shocks are incorporated into the prices of the most liquid agricultural commodity futures, namely those traded at the Chicago Mercantile Exchange. The distributed nature of agricultural commodity production leads to some natural questions regarding the formation of futures prices at the CME. Do supply shocks at locations far from the US Midwest matter as much as those close to Chicago? How is the size of local production related to the response of

the CME price to local events? How is the rise of emerging economies in the global production of commodities represented in the dynamics of CME prices?

CME contracts and prices are used and influenced by a wide variety of traders across the world. This includes participants in the soybean production and intermediation chain in the US, and also economic agents in exporting and importing nations in Latin America and Asia. Regunaga (2010) noted that trading and crushing companies in Argentina, which include local offices of large transnational firms such as Bunge, Cargill and Glencore as well as domestic firms, hedge their price risk with CME contracts. Christofolletti et al. (2012) found evidence of increasing strength in the link between CME soybean prices and prices of soybean contracts at the Dalian exchange in China. Therefore, beyond their academic interest, answers to the questions explored in this paper are relevant for policy makers and risk managers. I find that CME soybean futures prices have become increasingly sensitive to supply shocks outside of the US. This implies that firms that hedge their sales and purchases by trading CME contracts are doing so at price levels that are increasingly dependent on globally distributed supply shocks.

I explore the broad questions outlined above by focusing on the global soybeans market, which is particularly convenient for three reasons. First, local summer rain precipitation tends to increase local soybean output. Therefore rain is an exogenous source of local supply shocks. Second, the US Midwest and the central region of Argentina concentrate about 45% of the global production of soybeans. Spatial concentration is convenient because it amplifies the impact of local events on global prices¹. Finally, the market shares of these dominant regions have varied over the last fifteen years due to strong output growth in Argentina and much slower growth in the US. This allows me to track how the response of the CME price to local supply shocks has evolved following

¹Output from Brazil, the second largest producer of soybeans, is taken into account in the aggregation of global supply, but rain in Brazil is not used a source of supply shocks for reasons discussed in Section 2.

changes in market shares. In summary, the combination of these three elements allows me to identify how the price of soybeans at the CME responds to regional supply shocks and to relate the price impact of rain to shares of global production.

Soybeans are grown from late spring to early fall, so I define the *extended summer* as May-October for the US Midwest and November-April (of the following year) for Argentina. Output at harvesting time is most sensitive to rain during the *midsummer* months, defined as July and August for the US Midwest, and January and February for Argentina. I find that in the period 1996/7 to 2002/3 one inch of midsummer rain over the US Midwest led, on average, to a 5.2% decrease in the CME soybean price after controlling for changes in the Oil price and local temperature among other factors. During that period, one inch of midsummer rain over the central region of Argentina had no significant impact on the CME price. In this period, the production of soybeans in the US Midwest accounted for 30.8% of global annual production, and the share of global output grown in the central region of Argentina was only 12.3%. In the more recent period of 2003/4 to 2010, coincident with a decrease of the US Midwest share to 23.6% and an increase of Argentina's central region's share to 16.5%, I find that one inch of midsummer rain over the US Midwest led to a 1.8% decrease in the relevant CME soybean price while the same amount of rain in Argentina led to a 1.5% CME price decrease. For proper context, midsummer rain precipitation in any of these two regions has historically been close to 10 inches. Therefore, one inch of rain is biologically important and leads to the economically significant price impact that I estimate. Analogous estimates of rain's impact over a partition of the 1996-2010 dataset on biennial periods shows that the magnitude of local rain's impact on price, across periods and regions, was close to linearly related to the corresponding local share of global production.

I show that these empirical findings are generally in agreement with the predictions of a very simple model of distributed commodity production that assumes that an ad-

ditional unit of supply anywhere in the world has the same impact on the CME price. My empirical results suggest that traders at the CME incorporate information about distributed supply shocks in a manner that is roughly consistent with a globally integrated market for soybeans and indifferent to the geographical origin of supply. Therefore, as output from Latin America has recently grown relative to that from the US, risk managers using CME contracts are hedging their output at prices increasingly influenced by events outside of the US.

The setting chosen in this paper to explore the spatial aggregation of supply shocks into a single asset price is economically important. The growing weight of developing economies in the global economy is one the main developments of the last decade. Between 2002 and 2009, the GDP of developed economies including, among others, the Euro area, Japan, the UK and the US, grew at an average annual rate of 1.3%. During the same period, developing economies including China, India, Argentina and Brazil, achieved an annual rate of growth of 5.8% (United Nations' *World Economic Situation and Prospects 2011* ²). Some of the most important developing economies are heavily dependent on commodity markets. Argentina and Brazil are leading exporters of agricultural commodities including soybeans, corn, wheat, cotton and coffee. Brazil is also the third global producer of iron ore, behind China and Australia. Russia is the second world producer of oil and it has a large presence in the metals market. China is a leading commodity importer and is expected to grow its influence in the near future.

Commodities are traded globally through financial and physical contracts. For historical and institutional reasons the most important derivatives market for agricultural commodities is the Chicago Mercantile Exchange (CME), which merged with the Chicago Board of Trade (CBOT) in 2007. Futures contracts for soybean and derived products traded at the CME are very liquid. CME quoted prices are widely followed by soybean

²<http://www.un.org/en/development/desa/policy/wesp/>

farmers, traders, hedgers and speculators around the world. CME contracts are also used by investors, specialized hedge funds and index funds increasingly interested in gaining exposure to commodity markets as an asset class (Tang and Xiong (2012)).

This paper is related to earlier work in the literature. The question of how futures prices reflect fundamental information has been central in the asset pricing literature in general and commodities in particular. Within this line of work, this paper is closest in spirit to a few studies linking weather related information and asset prices. Roll (1984) was among the first to study the effect of weather in agricultural futures prices by evaluating the impact of rain and temperature on the production of orange juice in Florida. He found surprising little explanatory power for meteorological variables. His conclusions were then revised by Boudoukh et al. (2007), who found strong explanatory power for a nonlinear effect of temperature on orange juice prices due to negative supply shocks caused by freezing events. As it is the case for soybeans, oranges for juice production are heavily concentrated in Florida and the state of Sao Paulo in Brazil. Boudoukh et al. (2007) provided some anecdotal evidence about the importance of news about Brazilian production reported in the Wall Street Journal, and performed a correlation analysis between US and Brazilian production and exports based on annual data. They found evidence that Brazilian and Florida oranges are substitutes but did not perform a direct analysis of Brazilian supply shocks on US prices. In contrast, this paper analyzes the behavior of CME prices using an econometric methodology that treats the US and Argentina symmetrically and relies on high frequency local weather data. Fleming et al. (2006) studied the interplay between the flow of information originated in weather dynamics and changes in commodity prices, including soybeans, by analyzing variance ratios for weather sensitive markets. The impact of changes in expected supply and demand into agricultural commodity prices was studied by Isengildina-Massa et al. (2008) who estimated the response of CME soybean futures prices to the widely followed World

Agricultural Supply and Demand Estimates (WASDE) released monthly by the US Department of Agriculture. They found a significant increase in price variance on the dates of a WASDE release. Their work, however, is silent about the mechanism of aggregation of supply shocks from different geographical locations. Moreover, the WASDE report is computed by the USDA using fundamental sources of information that might also be processed independently by sophisticated market participants and incorporated into prices, to some extent, prior to the release of the report. In this paper I use rain precipitation that is truly unanticipated information beyond the horizon of weather forecastability. Second, the WASDE report is released monthly, while rain precipitation is available with daily frequency, potentially allowing for more precise estimates. Other papers studying the economic impact of weather events on agricultural activity include Roberts and Schlenker (2010). They exploited exogenous weather related supply shocks to estimate demand and supply elasticities of agricultural commodities and used their estimates to evaluate the impact of biofuels on food prices.

Work on cross-market information flow in commodities has explored the notion that the same physical asset should be priced consistently at different locations after taking into account transportation costs, tariffs, and other frictions. Yiang et al. (2000) studied the validity of the law of one price for soybean, focusing on the US, Argentina and Brazil. Findlay and O'Rourke (2003) analyzed major trends on commodity market integration across a 500 year span. Fung et al. (2010) studied the flow of information in aluminum and copper markets between the New York Mercantile Exchange and the Shanghai Futures Exchange. Chng (2009) studied the joint dynamics of seemingly unrelated commodities used by the same industry. Christofolletti et al. (2012) studied the price linkages between soybean prices in Argentina, Brazil, China, and the US. In this paper, rather than asking how commodity prices and returns are aligned across markets, I ask how local supply shocks information is incorporated into the CME price that serves

as a global benchmark, paying special attention to the relevance of regional shares of global production.

The paper is structured as follows: in Section 2, I describe the structure of the global soybean market. In Section 3, I present a simple model of a globally integrated economy in which regional market shares are linearly related to the sensitivity of the CME price to regional rain. This model is estimated using the econometric approach described in Section 4, applied to the data presented in Section 5. My empirical results are in Section 6 and conclusions in Section 7.

2. The Global Soybean Market

Soybeans and derived products are very important protein sources of agricultural origin. The so-called soybean complex includes soybeans, soymeal primarily used for animal feeding, and soybean oil. Crushing a ton of soybeans yields approximately 0.73 tons of meal and 0.18 tons of oil. The soybean market has grown very strongly over the last twenty years. Annual production in the World, the US, Argentina, and Brazil during the period 1996-2010, measured in millions of metric tons per year, and the corresponding shares of global production are reported in Table I. In 1997 Argentina grew 8% of global production, and this became 19% in 2011. During the same period, the relative contribution of the US production to global output fell from 49% to 35%. Table I also shows that yields have been similar in Argentina and the US and have remained relatively stable in time.

Table II shows soybean and soymeal exports from various regions. Argentina exports the bulk of its production in the form of soymeal and soybean oil and only a small fraction of its production in the form of beans. The US exports mostly in the form of beans. Domestic consumption of soybean related products is significant in the US and

very small in Argentina. Tables I and II are composed using information from the USDA *Oilseeds Reports* for December 2006, December 2010 and May 2011 ³. The price of a metric ton of soybeans at the CME on August 2012 was approximately 650 US dollars. For a global production of 260 million metric tons this implies a market value of roughly USD 170 billion.

Argentina and Brazil, the second largest soybean producer after the US, experienced in recent decades similar rates of growth. I focus in this paper on the impact of rain in the central region of Argentina and the US Midwest because these soybean producing regions are biologically similar in the impact of precipitation on soybean yields. The main soybean producing regions in Brazil are the states of Goias, Matto Grosso, Parana, and the northern part of the state of Rio Grande do Sul, which have historically received summer rainfall exceeding that in Argentina and the US Midwest by approximately 50%, therefore making its output potentially less sensitive to rain. More importantly, the precipitation data used in this paper for Argentina and the US Midwest was gathered from the database at the National Climatic Data Center (NCDC) administered by the National Oceanic and Atmosphere Administration (NOAA). This source, which reproduces information submitted by local official meteorological organizations, exhibited severe gaps between 2000 and 2004 for Brazilian weather stations. Although Argentina and Brazil are neighboring countries in South America, the main soybean producing regions in Brazil are between 500 and 1500 miles away from the soybean producing region in central Argentina, therefore affected by different high frequency weather events.

Table III reports annual soybean production of individual provinces in Argentina. The most important soybean producing region in Argentina is conformed by the provinces of Buenos Aires, Santa Fe, Cordoba and Entre Rios, located at the center of the country. Between 1997 and 2003 this region produced an average of 89% of national output, a ratio

³http://www.fas.usda.gov/oilseeds_arc.asp.

that decreased slightly to 87% for the period 2004-2010⁴. Table III also shows annual regional yields (in tons per hectare). Table IV shows annual output for states in the US Midwest, defined as the region including Illinois, Indiana, Iowa, Missouri, Minnesota, Nebraska, and Ohio. The US Midwest produced 72% of US soybeans during 1996-2002 and 68% of US soybeans during 2003-2010. Table IV also reports annual regional yields.

Soybeans are grown especially during the summer and more generally from late spring to early fall. This ranges from May to October in the northern hemisphere, and from November to April in the southern hemisphere. Soybean supply at harvesting time is determined by the sum of production in the current season and small existing stocks from previous years. Figure 1 shows the time-line of production and harvesting events in Latin America and the US Midwest. Soybeans are usually consumed or processed in the year following harvest. The highest ratio of existing stocks in the US at the beginning of the year to subsequent yearly production for the period 1997-2010 was 22%, and less than 10% for most years. The same ratio for Argentina and Brazil in the same period of time was never above 18% and also less than 10% for most years. (Tables 21, 22 and 23 in the *Oilseeds Report* of December 2010, USDA). Therefore local supply at year end is for the most part provided by newly harvested soybeans.

Production is determined by two factors: the area of land allocated each year to growing soybeans, and the yield per acre which measures output per unit of land. The year to year variation in the area of land allocated to soybeans has an impact on supply but it is a low frequency variable because it is determined only once per year and the estimation of its effect on prices is complicated by its endogeneity. Yield depends on the available growing technology, quality of the soil, use of fertilizers and herbicides and, most crucially, water intake. Irrigation technology is expensive and not used in the

⁴Production information at the province level for the 2011 harvesting season was not available at the time of writing this paper

soybean growing regions in Argentina. Irrigation in Midwest states reported by the US Geological Survey in (Hutson et al.(2004)) shows that the proportion of irrigated land for growing soybeans is negligible for every state in the US Midwest, perhaps with the exception of Nebraska. I could not determine how much of the irrigated land in Nebraska is allocated to growing soybeans, but the fact that irrigation there covers only 15% of the state (Hutson et al.(2004)) and that Nebraska has historically produced about 10% of Midwest soybean output leads me to assume in the remainder of the paper that water intake for growing soybeans in the US Midwest is essentially provided by rain.

Tannura et al. (2008) identified a significant and positive correlation between summer rain precipitation and soybean yields in the states of Iowa, Illinois and Indiana. The association was found to be particularly strong for the months of July and August, which is the period I call midsummer. I provide in this paper additional empirical support for the strong link between midsummer precipitation and soybean output for Argentina and the US Midwest. This biological link between exogenous rain and output will be exploited in this paper to identify the sensitivity of soybean prices to supply shocks. From an econometric perspective, the availability of daily recordings for a large number of spatially distributed weather stations, leading to many data points per season, makes rain a very convenient variable to study the impact of exogenous changes in expected supply on the price of soybeans. Tannura et al. (2008) also found that temperature fluctuations lead to variability in output in the US Midwest, therefore in my econometric analysis I also include temperature as a potential source of price variations.

Commodities are traded physically by producers, consumers and big trading houses such as Cargill, Bunge, Noble and Glencore. Physical transactions are associated to a variety of shipping routes across the globe. Freight-on-board (FOB) prices for spot physical delivery at certain shipping terminals are quoted on a daily basis in the over-the-counter market for the Gulf of Mexico and several ports in Latin America. These

FOB prices are usually positively correlated with the CME price but differ from it by a stochastic basis driven by seasonal and technical factors. Hedging and speculation on commodities is mediated by financial contracts. The prices of the most liquid contracts for soybeans, traded at the CME, are widely followed by soybean traders around the world including traders based in Argentina as discussed by Regunaga (2010). Soybeans underlying CME contracts are delivered at expiration in physical form at warehouses in the US Midwest, although most futures contracts are canceled before expiration.

3. The Model

The main contribution of this paper is the family of empirical results in Section 6, which explicitly link regional rain with CME price changes, and rain's impact on prices with global market shares of production. Yet, in order to understand the significance of these empirical results I find it helpful to introduce a highly stylized model of distributed production under exogenous weather shocks. In the remainder of this Section I will focus explicitly on the effect of rain on prices, although temperature fluctuations will receive equal attention in the econometric tests of the model.

3.1. Rain and supply

A large number of firms produce soybeans in Argentina, Brazil and the US and sell their output in a competitive global market. Between 1996 and 2010, these three countries had close to an 80% market share, as seen in Table I. Deodhar and Sheldon (1997) studied the degree of competition in the soymeal export market using data with annual frequency from 1966 to 1993 and found that the market was very close to perfectly competitive. Beans are grown during the extended summer season lasting from the beginning of May to the end of October in the US, and from the beginning of November to the end of April

in Latin America.

This is a model for a single growing season (six months long). Let T be the harvesting time at the end of the current growing season, measured in days. Because Latin America and the US are in opposite hemispheres, only one of these two regions is actively growing soybeans. For the active region, total local supply at T ,

$$Q_{Supply,T}^{ActiveRegion} \quad (1)$$

as seen from $t < T$ is uncertain because production will depend on weather events.

The supply from inactive regions at T is the inventory that remains from earlier harvests (the last one being roughly six months in the past). This is inventory available in April in the US, or inventory available in Latin America in October. Figure 1 represents the chronology of the main events. Consider, to fix ideas, the growing season from month 0 to month 6 in Figure 1. In this case the active region is the US, while Latin America is inactive. I assume that the rate of soybean consumption within a growing season is known, therefore inventory at the inactive region is predictable as seen from $t < T$. In particular, the ratio of existing stocks in the US, Brazil and Argentina at the beginning of the local year to subsequent yearly production for the period 1997-2010 was most often less than 10%, and never more than 20% (Tables 21, 22 and 23 in the *Oilseeds Report* of December 2010, USDA). Taking into account this small $\approx 10\%$ initial inventory prior to the last harvest, a plausible linear consumption rule implies that supply at the inactive region roughly 6 months after the last harvest is

$$Q_{Supply,T}^{InactiveRegion} \approx Q_0^{InactiveRegion} 0.6 \quad (2)$$

where Q_0^{Region} is the long term average regional yearly output. Total global supply available at T is

$$Q_{Supply,T}^{World} = Q_{Supply,T}^{US} + Q_{Supply,T}^{Argentina} + Q_{Supply,T}^{Brazil} + Q_{Supply,T}^{Rest}. \quad (3)$$

Total expected supply at T , conditional on the information I_0 available at $t = 0$, can be decomposed in the contributions from active and inactive regions:

$$E[Q_{Supply,T}^{World} | I_0] = Q_0^{Active} 1.1 + Q_0^{Inactive} 0.6 = Q_0^{World} (S^{Active} 1.1 + S^{Inactive} 0.6)$$

where S^{Region} is the market share of the region of interest in the global production of soybeans. Market shares are reported in Table I. By alternatively assigning the role of active and inactive region to the US, and the combined output of Argentina and Brazil, and splitting $Q_{Supply,T}^{Rest}$ equally on active and inactive regions, it can be seen that in the period 1996-2010 it held that

$$0.8 < S^{Active} 1.1 + S^{Inactive} 0.6 < 0.9$$

therefore

$$E[Q_{Supply,T}^{World} | I_0] \approx Q_0^{World} 0.85 \quad (4)$$

Hence, total relative supply available at T , as a fraction of yearly global production, was relatively more stable than the individual outputs of the US or Argentina over the same period. This is due to fact that the US and Latin America had comparable outputs in this period, and the mitigating effect of inventories. It is important to note that output from Brazil is appropriately included in the present analysis for the accounting of the *level* of global supply. However, in this paper I will focus on supply *shocks* in Argentina and the US Midwest, for reasons already discussed in Section 2.

In the active region, uncertain output at T depends on future production. The amount of land, A_0^{Region} , allocated to growing soybeans is set at $t = 0$. The terminal yield obtained

at harvesting time in the active region is most sensitive to rain during the midsummer months of July and August in the US and January and February in Latin America. Regional output is harvested at T in the region that is currently active and added to typically small existing stocks, for final regional supply $Q_{Supply,T}^{Region}$ measured in metric tons. Land, terminal yield Y_T^{Region} , and supply are related through

$$Q_{Supply,T}^{Region} = Q_0^{Region} 0.1 + A_0^{Region} Y_T^{Region}. \quad (5)$$

Let T_{start} and T_{end} be the start and the end of the relevant rain period within a growing season, so that $0 < T_{start} < T_{end} < T$. On the i -th trading day, $i = T_{start}, \dots, T_{end}$ let I_i be the information available to CME market participants. Expected regional soybean supply at harvesting time T , conditional on the available information, is $E[Q_{Supply,T}^{Region} | I_i]$.

Uncertainty about supply at harvesting time is a function of the yield per unit of land, which depends on technology and water intake. In the absence of irrigation, water intake is provided exclusively by rain. Let R_j^{Region} be the amount of rain fallen on day j over a region of interest and

$$R^{Region} = \sum_{j=T_{start}}^{T_{end}} R_j^{Region} \quad (6)$$

the total precipitation during the midsummer period. Based on the positive relationship between precipitation and yields documented by Tannura et al. (2008) I postulate a linear relationship between midsummer rain and terminal output at harvesting time

$$Q_{Supply,T}^{Region} = Q_0^{Region} 0.1 + Q_0^{Region} (1 + \gamma^{Region} (R^{Region} - \bar{R}^{Region})) \quad (7)$$

where \bar{R}^{Region} is the long term average midsummer rain and the positive coefficient γ^{Region} is the relative sensitivity of regional yield to total rain during the midsummer. Because the amount of land allocated to soybeans is fixed within a growing season, the rain dependent term in (7) is equivalent to

$$Y_T^{Region} = \bar{Y}^{Region} + \gamma^{Region} \bar{Y}^{Region} (R^{Region} - \bar{R}^{Region}) \quad (8)$$

where \bar{Y}^{Region} is the long term historical yield. Expectations of terminal output are updated during the midsummer reflecting changes in expected rain precipitation. Taking conditional expectations in (7) I write

$$E[Q_{Supply,T}^{Region}|I_i] - E[Q_{Supply,T}^{Region}|I_{i-1}] = \gamma^{Region} Q_0^{Region} (E[R^{Region}|I_i] - E[R^{Region}|I_{i-1}]). \quad (9)$$

A precipitation forecasting technology such as that provided by the National Weather Service is available to market participants on each day. If this technology, conditional on I_i , has perfect predictive capability for up to τ days beyond $t = i$ and it is completely unskilled beyond that horizon, then I can write

$$E[R^{Region}|I_i] = \sum_{j=T_{start}}^{i+\tau} R_j^{Region} + ((T_{end} - T_{start}) - i - \tau) \frac{\bar{R}^{Region}}{T_{end} - T_{start}} \quad (10)$$

because daily rain that is beyond τ days into the future is unpredictable therefore its expectation is $\frac{\bar{R}^{Region}}{T_{end}-T_{start}}$, and rain that has already occurred, or that will occur in a horizon shorter than τ has already been incorporated with certainty into the expectation of future precipitation. Therefore,

$$E[R^{Region}|I_i] - E[R^{Region}|I_{i-}] = R_{i+\tau}^{Region} - \frac{\bar{R}^{Region}}{T_{end} - T_{start}}. \quad (11)$$

It is learning about rain at the prediction horizon that drives the daily change in expectations in (11). In a more realistic setting, weather forecasting skill is less than perfect and certainty about rain on a given target date develops gradually as time evolves towards the target date. Bickel et al. (2011) evaluated the accuracy of Probability of Precipitation (PoP) forecasts generated by the National Weather Service dependent from NOAA, the Weather Channel, and CustomWeather, which are available to market participants at low or no cost. They considered forecasts at 753 location in the US and found evidence of significant forecasting skill for up to a 96 hour horizon for all weather forecast providers. This leads me to postulate

$$E[R^{Region}|I_i] - E[R^{Region}|I_{i-1}] = \frac{1}{\tau} \sum_{j=1}^{\tau} R_{i+j}^{Region} - \frac{\bar{R}^{Region}}{T_{end} - T_{start}} + \eta, \quad (12)$$

where η is a generic random error uncorrelated with $\sum_{j=1}^{\tau} R_{i+j}^{Region}$, and $\tau = 4$ days.

From (9) and (12) it follows that the impact of rain in expected supply is

$$E[Q_{Supply,T}^{Region}|I_i] - E[Q_{Supply,T}^{Region}|I_{i-1}] = \gamma^{Region} Q_0^{Region} \left(\frac{1}{\tau} \sum_{j=1}^{\tau} R_{i+j}^{Region} - \frac{\bar{R}^{Region}}{T_{end} - T_{start}} \right) + \eta \quad (13)$$

for a rescaled error term η . An alternative to the approach described above consists in using actual weather forecasts rather than subsequent precipitation. However, there is no publicly available record of historical weather forecasts for Argentina going back as

far as 1996. Regarding US weather forecasts, Chincarini (2011) reports that archived forecasts for the National Weather Service Model Output Statistics begin in September 2005.

3.2. *Supply and prices*

Let P_T be the soybean spot price at the end of the current growing season and let $Q_{Demand,T}^{World}$ be global demand at T . I adopt a standard constant elasticity demand curve

$$Q_{Demand,T}^{World} = aP_T^b, \quad (14)$$

with price elasticity of demand $b < 0$, and assume that this remains unaffected by local regional rain during the growing season. This is plausible because soybeans are used primarily for the delayed production of soymeal and soybean oil and sold for later consumption at widely spread locations (often overseas) that are far from the relatively small producing regions. In addition, the fact that several months are needed to grow soybeans leads me to postulate that $Q_{Supply,T}^{World}$ is inelastic in the short run with respect to prices. From (14), equilibrium at T implies

$$aP_T^b = Q_{Supply,T}^{World}. \quad (15)$$

World supply in (15) is the sum of supply at T in active and inactive regions. The latter contribute to total supply but are assumed to be certain because they are remaining inventories from earlier harvests. As such, the day to day variation in expected supply at T is driven by changes in expected supply in the active regions. Taking conditional expectations with respect to the information available on consecutive days in the growing season, summing over active regions, and using (13) I obtain

$$E[aP_T^b|I_i] - E[aP_T^b|I_{i-1}] = \sum_{ActiveRegions} \gamma^{Region} Q_0^{Region} \left(\frac{1}{\tau} \sum_{j=1}^{\tau} R_{i+j}^{Region} - \frac{\bar{R}^{Region}}{T_{end} - T_{start}} \right) + \eta^{Region}. \quad (16)$$

Next, let $P_0 \equiv E[P_T|I_0]$. I assume that changes in expected price within a growing season are relatively small so that

$$|P_T - P_0| \ll P_0 \quad (17)$$

Under this assumption it is valid to make a linear expansion of P_T^b , as a function of P_T , in the vicinity of P_0 ,

$$P_T^b \approx P_0^b + bP_0^{b-1}(P_T - P_0) \quad (18)$$

Next, take conditional expectation with respect to I_0 on (15) and use (18) to obtain

$$E[Q_{Supply,T}^{World}|I_0] = E[aP_T^b|I_0] \approx aP_0^b. \quad (19)$$

Also, taking conditional expectation in (18) with respect to information on consecutive days leads to

$$E[aP_T^b|I_i] - E[aP_T^b|I_{i-1}] \approx abP_0^{b-1}(E[P_T|I_i] - E[P_T|I_{i-1}]). \quad (20)$$

Therefore using (20) in (16) implies

$$\frac{abP_0^b(E[P_T|I_i] - E[P_T|I_{i-1}])}{P_0} \approx \sum_{ActiveRegions} \gamma^{Region} Q_0^{Region} \left(\frac{1}{\tau} \sum_{j=1}^{\tau} R_{i+j}^{Region} - \frac{\bar{R}^{Region}}{T_{end} - T_{start}} \right) + \eta^{Region}$$

and then, by (17) and (19),

$$\frac{E[P_T|I_i] - E[P_T|I_{i-1}]}{E[P_T|I_{i-1}]} \approx \frac{1}{bE[Q_{Supply,T}^{World}|I_0]} \sum_{ActiveRegions} \gamma^{Region} Q_0^{Region} \left(\frac{1}{\tau} \sum_{j=1}^{\tau} R_{i+j}^{Region} - \frac{\bar{R}^{Region}}{T_{end} - T_{start}} \right) + \eta^{Region} \quad (21)$$

which, by (4) is

$$\frac{E[P_T|I_i] - E[P_T|I_{i-1}]}{E[P_T|I_{i-1}]} \approx \sum_{ActiveRegions} \frac{\gamma^{Region} S^{Region}}{0.85b} \left(\frac{1}{\tau} \sum_{j=1}^{\tau} R_{i+j}^{Region} - \frac{\bar{R}^{Region}}{T_{end} - T_{start}} \right) + \eta^{Region}. \quad (22)$$

where $S^{Region} \equiv \frac{Q_0^{Region}}{Q_0^{World}}$ is the market share of the region of interest in the global production of soybeans and dt is a deterministic trend.

The analysis so far has related expected spot price at harvesting time T with expected future rain. I now make the connection with the CME futures price available at $t = i$. Let P_i^T be the CME soybean future price observable at $i \leq T$ associated with a CME contract expiring at the end of the current growing season. Let the future price be the expected spot price at expiration plus a small risk premia Y_i

$$P_i^T = E[P_T|I_i] + Y_i. \quad (23)$$

Then, (22) can be expressed using (23) to obtain

$$\frac{P_i^T - P_{i-1}^T}{P_{i-1}^T} \approx \sum_{ActiveRegions} \frac{\gamma^{Region} S^{Region}}{0.85b} \left(\frac{1}{\tau} \sum_{j=1}^{\tau} R_{i+j}^{Region} - \frac{\bar{R}^{Region}}{T_{end} - T_{start}} \right) + \eta^{Region} + dt. \quad (24)$$

This is a highly stylized model for a fully integrated global market in which an additional unit of supply anywhere in the world has equal impact on the CME price traded in Chicago. Supply shocks, in turn, depend on regional rain, market shares, and physical characteristics of each region.

The first quantitative prediction of the model (24) is that excess local rain precipitation leads to a decrease in the CME future price through the combination of the price elasticity of demand $b < 0$ and the regional yield sensitivity with respect to rain $\gamma^{Region} > 0$.

In this model, rain in the active region has a positive impact on supply at T , and therefore a negative impact on the price for delivery then. As represented in Figure 1, the same consumption/inventory holding rule in (2), implies that

$$Q_{Supply, T+6months}^{ActiveRegion} \approx Q_T^{ActiveRegion} 0.6. \quad (25)$$

Therefore, rain in the active region between 0 and T also contributes, partially, to local supply on $T+6$ months. Global supply on $T+6$ months is the sum of inventory carried over from T , and output from the new harvest at $T+6$ months. It follows from the setup of the model that, as a consequence of the possibility of storage, rain in the currently active region should also cause a negative price shock for longer dated futures, but of decreasing magnitude as a function of maturity.

The third prediction of the model (24) is that the sensitivity of the CME price with respect to regional rain is proportional to $\gamma^{Region} S^{Region}$. This is a testable implication that follows directly from the assumption about global integration of supply. A special case of interest is that of physically similar regions such as the US Midwest and the central region of Argentina. If both regions have equal sensitivity γ^{Region} , then the response of the CME price to local rain is simply proportional to the local share of the global market. Rather intuitively, according to this model the CME price should be more responsive to weather events on those regions that produce larger quantities of soybeans.

The model assumes that a single variable measuring total rain over an entire region predicts regional output. This is clearly a simplification since rain and yields also exhibit

variability within each region. However, attempts to identify the price impact of rain on smaller regions (e.g. individual states) are likely to be hampered by high rain correlation between neighboring states and the lesser economic importance of each individual state. The partition in two very distinct regions adopted in this paper seems appropriate to investigate the implications of globally distributed production.

The model is silent about other factors that might affect soybean prices. Tannura et al. (2008) showed that soybean output is sensitive to summer temperatures therefore I add temperature fluctuations as a control in the estimation stage. Prices are also sensitive to demand fluctuations and financial speculation. These issues, although certainly very important in the determination of commodity prices, seem at first order unrelated to the main goal of this paper, namely evaluating the impact of the distributed nature of production. Some of these effects, formally included in the noise term η in this Section, are taken into account in the empirical estimation of the model through appropriate controls.

4. The Econometric Specification

A central premise in this paper, namely that regional rain has an impact on CME prices, relies on the fact that rain is a source of exogenous supply shocks because it tends to increase yield and therefore output, as documented by Tannura et al. (2008). I begin by testing the validity of this assumption for the sample period used in this paper. I estimate the response of regional yield to regional rain (8) by

$$Y_i^{Region} = \alpha^{Region} R_i^{Region} + const + \epsilon_i \quad (26)$$

using annual recordings of yield Y_i^{Region} and summer rain R_i^{Region} between 1997 and 2010

for the US Midwest and the central region of Argentina. The rain variable is constructed in this exercise as the sum of total precipitation over all weather stations in the region of interest during the midsummer or extended summer of a given year. Results of this estimation confirm that rain is indeed a source of supply shocks and, by comparing $\alpha^{Midwest}$ and $\alpha^{Argentina}$, support the notion that these two regions are similar in their yield response to rain.

Next, I run daily frequency estimations of the model (24) for Argentina and the US Midwest, on periods of time with significantly different regional market shares. This allows me to identify the impact of rain in each region, and how this impact changed as the relative weight of Argentina's production increased.

The daily rain variable for any of the two regions being considered is the average rain fallen at the weather stations over the entire region of interest during four consecutive days, including rain that occurred on weekends. For day i , the rain variable is

$$Rain_i^{Region} \equiv \frac{1}{4} \frac{1}{Nstations} \left(\sum_{j=1}^4 \sum_{k=1}^{Nstations} R_{i+j}^{stationk} \right), \quad (27)$$

where $Nstations$ is the number of stations in the region of interest. For every summer in each period I compute the daily price change and daily rain precipitation variable in every trading day between the beginning of January and the end of February, and estimate the theoretical return (24) for Argentina through the econometric specification⁵

$$\frac{P_i^{May} - P_{i-1}^{May}}{P_{i-1}^{May}} = \beta_{Arg} Rain_i^{Arg} + controls_i + Const + \epsilon_i, \quad (28)$$

The CME price P^{May} corresponds to the contract expiring on the forthcoming May, at

⁵Regressions for (28) and (29) are run on daily rain, daily price changes and daily control changes. I verified that all these variables are stationary in the sample under consideration.

the end of the current growing season as seen from day i . I estimate (24) for the US Midwest using the approach outlined above but using US Midwest rain precipitation data and daily price changes recorded from the beginning of July to the end of August,

$$\frac{P_i^{Nov} - P_{i-1}^{Nov}}{P_{i-1}^{Nov}} = \beta_{US} Rain_i^{US} + Controls_i + Const + \epsilon_i. \quad (29)$$

The price variable P^{Nov} corresponds to the CME contract with expiration at the end of the current growing season as seen from day i . The separate regressions that I run for Argentina and the US are symmetric in the fact that I use price changes for futures delivery dates at the end of the respective growing seasons.

The specifications (28) and (29) test for the impact of local rain on specific contracts. Statistically significant negative results in these estimations would indicate that CME prices respond as expected to positive local supply shocks. Although this would be in agreement with the first prediction of the model in Section 3, such a finding would not rule out a segmented market, with CME May contracts responding exclusively to Latin America's events, and CME November contracts responding solely to US shocks. In order to distinguish between this case, and that of an integrated soybean market with global inventory held across time, I structure appropriate tests. First, the impact of local rain on futures expiring beyond the local harvesting date is estimated by

$$\frac{P_i - P_{i-1}}{P_{i-1}} = \alpha + \beta_{Arg} Rain_i^{Arg} + \beta_{US} Rain_i^{US} + controls_i + \epsilon_i, \quad (30)$$

for the joint effect of rain precipitation in Argentina and the US on the daily return of CME Soybean future prices. Price changes are computed for a fixed May or November contract. In any specific regression, the sample for a fixed contract includes rain in both regions during the previous January, February, July and August (possibly from the

previous year), corresponding to two consecutive summers (one for Argentina and one for the US). By construction, this specification also includes winter rain, because midsummer months in the US coincide with midwinter months in Argentina, and viceversa. This specification allows me to test whether both contracts are simultaneously sensitive to rain in Argentina and the US and potentially rule out segmentation in the term structure of futures prices.

Finally, I test the third prediction of the model in Section 3., namely global market integration reflected in a linear relationship between price impact of rain and regional market shares for physically similar regions. As a preliminary step, (28), (29) and (30) are estimated on a partition of the data set in two long periods. For Argentina, the first period includes every day in the midsummers from 1997 to 2003, ending in February 2003, and the second period includes the midsummers from 2004 to 2010. The US Midwest first period includes the (northern hemisphere) midsummers, starting in July and ending in August, from 1996 to 2002. Midsummers from 2003 to 2010 belong in the second period. The data available from the beginning of 1996 up to the time of writing this paper includes fourteen summers for Argentina and fifteen summers for the US. The discrepancy between periods across regions is minimal given the fact that summers in these regions are six months apart. Soybean production from Latin America, including Argentina and Brazil, surpassed US output in 2003, therefore this event coincides roughly with the break between the two periods I consider. I also obtain more frequent estimates of β from (30) by using biennial periods between 1997 and 2010. I then use these estimates of β_{US} and β_{Arg} to run a second regression that explicitly tests for a linear relationship between market share and rain impact. This is

$$\beta_{Region,Period} = \delta S_{Region,Period} + Const + \epsilon_{Region,Period}, \quad (31)$$

where $S_{Region,Period}$ is the average market share of the central region of Argentina or the US Midwest over each biennial period.

The simple model introduced in Section 3 that led to (28), (29) and (30) explains price changes by the supply effect caused by rain, and random rain forecasting errors. According to Tannura et al. (2008), soybean output is also negatively affected by high summer temperatures. I control by the effect of temperature on prices by adding an explanatory variable defined, for day i , as

$$Temp_i^{Region} \equiv \frac{1}{4} \frac{1}{Nstations} \left(\sum_{j=1}^4 \sum_{k=1}^{Nstations} Temp_{i+j}^{stationk} \right), \quad (32)$$

where $Temp_j^{stationk}$ is the average temperature recorded on day j at weather station k after filtering the seasonal component. Therefore, the spatial and temporal averaging in $Temp_i^{Region}$ is identical to that in the construction of $Rain_i^{Region}$.

Price changes are also caused by a myriad of other factors, discussed for example by Frankel and Rose (2009). The overall level of commodity prices has been found to be strongly influenced by the level of global economic activity that is closely related to energy prices. Moreover, agrochemicals used in the production of soybeans, and the crushing process, are both highly dependent on energy prices. I take these effects into account by including daily percentage changes in the price of West Texas Intermediate Oil, in US dollars per barrel, as a control. Commodity prices, denominated in US dollars, are potentially sensitive to the strength of this currency. I control for this effect by including daily percentage changes on an equally weighted basket of exchange rates between the US dollar vs. the Euro and the Chinese Yuan, which represent some of the most important economic regions involved in soybean international trade. The potential effect of varying transportation costs in the price of soybeans is taken into account by

adding daily percentage changes in the Baltic Dry Index as a control. This is a widely followed indicator of the cost of bulk freight shipping along several maritime routes used in international trade. Soymeal, directly derived from soybeans, is used to feed livestock, therefore I use the CME hog price as control as well.

5. The Data

5.1. *Price data*

I obtained daily CME price data from Bloomberg for the period May 1st, 1996 to October 29th, 2010. Soybean future prices are available every trading day for contract expirations in January, March, May, July, August, September and November. Underlying in a soybean future contract are 5000 bushels (approximately equal to 136.091 metric tons) of soybean of certain quality. The last trading day is the business day prior to the 15th calendar day of the contract month. The price is expressed as US dollar cents per bushel. Soybean prices for the nearest delivery contract ranged between 160 and 600 US dollars per ton between 1996 and 2010.

5.2. *Rain and temperature data*

Rain precipitation and average temperature data, both with daily frequency between January 1st, 1996 and December 31st, 2010, were obtained from the National Climatic Data Center (NCDC) administered by the National Oceanic and Atmosphere Administration (NOAA). Rain data is reported as inches of rain fallen within a 24 hour period. Average daily temperature is expressed in degrees Fahrenheit. Data is available for over 9000 stations distributed worldwide. Recordings are made in each country by a local meteorological governmental institution and exchanged with NOAA under the World Meteorological Organization (WMO) Weather Watch Program. I identified a list of representative weather stations in the central soybean growing region in Argentina and in

the US Midwest. Tables V and VI display statistics for total midsummer rain and average midsummer temperature at the selected weather stations in Argentina and the US Midwest respectively. In rare occasions (less than 3% and 1.6% for rain data in Argentina and the US respectively, even less often for temperature), a weather variable for a specific date and station is missing. In this case, the spatial average (27) is computed over the reduced number of weather stations. Unreported regressions treating the missing rain observations as zeros lead to results that are very similar to those reported in this paper.

5.3. *Controls*

All controls have daily frequency. The spot price for West Texas Intermediate oil, in US dollars per barrel, was obtained from the US Energy Information Administration (EIA) and denoted as Oil. The second control, denoted as FXRate is constructed using the nominal exchange rates between the Euro and the US dollar and between the Chinese Yuan and the US dollar obtained from the database in the Pacific Exchange Rate Service at the Sauder School of Business of the University of British Columbia. The third control is the Baltic Dry Index, denoted by BDI, and available from Bloomberg. CME Hog prices for near term futures were obtained from Reuters Datastream.

5.4. *Production data*

Global and regional soybean production and export data were collected from the USDA *Oilseeds Report* for December 2006, December 2010, and May 2011 ⁶, the USDA *Agricultural Statistics Report* for 1997, 2000, 2003, 2006, 2009, and 2011 ⁷, and from the online Agricultural Information Database of Argentina's Ministry of Agriculture, Livestock, and Fisheries ⁸. Production statistics, including regional yields, are presented in Tables I, III

⁶http://www.fas.usda.gov/oilseeds_arc.asp.

⁷http://www.nass.usda.gov/Publications/Ag_Statistics/

⁸<http://www.siiia.gov.ar/index.php/series-por-tema/agricultura>

and IV in Section 2.

6. Results and Analysis

6.1. *Rain and output*

As anticipated in Section 4, I begin by verifying empirically that regional rain has a positive impact on regional yield. Estimation results of (26) are reported in Table VII. The regressions show that the separate impacts of midsummer and extended summer rain on annual yield are statistically significant and positive for both regions. In addition, the hypothesis that both regions have equal sensitivity to local rain can not be rejected for the midsummer or extended summer periods. It is in this sense that I state, in the remainder of the paper, that the US Midwest and the central region of Argentina are biologically similar in their response to local rain. Results in Table VII also show that, in both regions, yields are more sensitive to rain during the midsummer than during the extended summer.

6.2. *Midsummer rain on prices*

I estimate (28) and (29) by OLS regressions with heteroscedasticity and autocorrelation robust Newey-West standard errors. Results in Table VIII show that midsummer rain in Argentina during the period 1997-2003 had no significant impact on CME May prices, and rain during the period 2004-2010 led, on average, to a 1.5% CME price decrease per inch of rain, significant at the 1% level. This is consistent with the notion that additional rain leads to a price decrease through an increase in expected output. Rain impact in the US Midwest led to a 5.2% price fall in the November contract during the first period, significant at the 1% level, and a 1.8% price fall during the second period, significant at the 10% level. The Oil price explanatory variable had a positive and strongly significant coefficient in the second period, for both regions, reflecting increasing correlation across

commodities, as documented by Tang and Xiong (2012). The FX Rate and the Baltic Dry Index did not have a consistent contribution to the explanation of CME price changes. Surprisingly, daily temperatures departures from their periodical seasonal variation had no statistically significant effect on CME price changes. In addition, unreported experiments using the average temperature over the next ten (rather than four) days, or using first order time differences in future temperature, also failed to find a discernible impact of these variables on CME price fluctuations. Hog prices did not appear to be significant across periods and regions.

In order to understand the economic significance of the estimated rain impact is useful to recall, from Tables V and VI, that historical average rain precipitation has been close to 10 inches for the midsummer in any of these two regions. Therefore, one inch of rain is biologically important and this precipitation leads to an economically significant price decrease in any of the two regions in the second period.

The results above, obtained by estimating (28) and (29), are convenient because they isolate the effect of rain in each specific region on contracts at the end of corresponding summers. However, in order to rule out market segmentation across contracts it is appropriate to estimate (30) for the joint effect of local rain in each contract. Results in Table IX show that midsummer rain in the US Midwest between 1997 and 2003 had a strongly significant negative effect in both May and November contracts. The impact of an inch rain, in any of these two contracts, was close to -3% . In the second period the impact of US rain was milder in significance and very similar across contracts. Rain in Argentina did not have a significant effect on CME contracts in the first period, and it had weakly significant negative effect in both contracts in the second period. The estimates faintly suggest that closely dated contracts (May for Argentina, November for the US) are more sensitive to local rain than longer dated contracts (November for Argentina, May for the US), in agreement with a decreasing impact of local supply shocks on future

inventories. The presence of winter rain in the sample for the estimation of rain impact coefficients in Table IX is likely to be the reason why these estimates are smaller in magnitude than those in Table VIII. In summary, the evidence suggests that the impact of local rain on the CME price is not contract specific.

6.3. Extended summer rain on prices

I evaluate the sensitivity of the results in Table VIII to the choice of summer months by estimating (28) and (29) using every month in the extended summer season rather than midsummer months. Table X shows rain's impact on the CME price for the central region of Argentina and the US Midwest. For the period 1997-2003 the impact of rain in Argentina was not significant, and for the period 2004-2010 the impact of rain was significant at the 5% level and led to a 0.6% fall in the CME price per inch of rain in the central region of Argentina. Results for the US Midwest in Table X show that for the period 1996-2002 the impact of rain in the US Midwest was significant at the 1% level and implied a 1.9% fall in the CME price per inch of rain in the US Midwest. This effect decreased in importance in the period 2003-2010, when the rain sensitivity in the US Midwest implied a 0.9% fall in the CME price per inch of local rain, significant at the 5% level. Oil and FXRate were strongly significant in the second period, with the expected positive and negative sign respectively.

A comparison between Tables VIII and X shows that the effect of an inch of rain on CME prices during midsummer months is much stronger than the effect during the extended summer months. This is consistent with the results reported earlier in Table VII which show that output is most sensitive to midsummer rain. The response of CME prices to rain during these months suggests that market participants are aware of this biological feature.

6.4. *Rain's price impact and market share*

The preceding empirical results show, through the statistical and economic significance of the OLS rain coefficients in Tables VIII, IX and X, that the CME market incorporates information about supply shocks caused by rain in the spatially and temporally distributed production of soybeans. Prices respond, during the midsummer or extended summer months, to rain on soybean producing regions on near future days, consistently with available weather forecasting technology. May and November CME contracts are sensitive to rain in both regions.

The main goal of this paper is to understand the relevance of distributed production in the global price formation mechanism. The very simple model in Section 3, which assumes that supply shocks are aggregated globally regardless of their origin, makes a specific prediction in this regard: the price impact of rain on a given region should be proportional to the product of the local yield sensitivity to rain γ^{Region} and the local share of global production S^{Region} . A special case of interest is that in which the hypothesis of γ^{Region} being equal across regions cannot be rejected. This is the case for the central region of Argentina and the US Midwest, and the model in Section 3 then predicts that price impact should simply be linear in global market share. This is the alternative considered first.

Multiplying the ratios representing the share of global output produced by Argentina reported in Table I, and the share of national output produced by the central region of Argentina reported in Table III, I find that the central region of Argentina produced an average of 12.3% of global output between 1997 and 2003 and 16.5% of global output between 2004 and 2010. This output growth coincides in time with the increased impact of Argentina's rain on CME prices documented in Tables VIII, IX and X. Identical computation, using the information in Tables I and IV implies global market shares for

the US Midwest of 30.8% and 23.6% for the first and second periods respectively. Not surprisingly, Tables VIII, IX and X show a decreased sensitivity of CME prices to US rain in the second period.

I test for the existence of a linear relationship between market share and price impact, consistent with $\gamma^{Argentina} = \gamma^{Midwest}$, by estimating (31). In order to generate a sufficiently large number of rain impact measurements across time I repeat the estimation of (30) for May and November contracts breaking the dataset in biennial periods, hence generating 14 point estimates of the impact of rain on daily price changes. Results are reported in Tables XI and XII. Although noisier than those reported in Tables VIII and X, all significant estimates in Tables XI and XII have negative sign as expected.

Market shares over biennial periods are computed using the information in Tables I, III, and IV in the manner described above. Market shares are reported in the first column of Table XIII. Rain impact estimates taken from Tables XI and XII are reproduced in the columns β_{May} and β_{Nov} in Table XIII. This information is also displayed in Figures 2 and 3, which make apparent a linear relationship between market share and rain's impact on the CME price. The statistical significance of this finding for the May and November contracts is reported in bottom panel of Table XIII. The slope coefficient is significant at the 1% level, and R^2 is in excess of 0.55. This is evidence of a strong empirical linear relationship between market share and local rain impact on the CME price. Such linear relationship is consistent with the prediction of the very simple model of Section 3 for regions that are biologically similar in their response to rain. Remarkably, CME traders seem to aggregate supply shocks linearly, regardless of their geographical origin.

An alternative to testing the model in Section 3 under the restriction $\gamma^{Argentina} = \gamma^{Midwest}$ as above, is to consider the unrestricted case in which regions are allowed to be different in their output response to rain. In this case, the model in Section 3 predicts, through (24), that regional market share should be proportional to the price impact of

rain β_{Region} divided by γ^{Region} . In turn, a comparison of (8) and (26), shows that α^{Region} in Table VII is an estimate of $\gamma^{Region}\bar{Y}^{Region}$. Therefore, in the unrestricted case, the model in Section 3 predicts that regional market share should be proportional to the ratio $(\beta_{Region}\bar{Y}^{Region})/\alpha^{Region}$. Renormalized rain impact coefficients are also presented in Table XIII, together with the results of a linear regression of the form (31) using the renormalized rain impact as dependent variable. Results in Table XIII show strong support for the predicted linear relationship. Therefore, the empirical evidence is consistent with a model in which CME traders aggregate global supply in a linear fashion, but it can not distinguish between the two alternatives just considered, namely if traders assume the regions are equal in their biological response to rain, or slightly different as suggested by the point estimates in Table VII.

Results in Table XIII, for both alternatives considered, also indicate the existence of a positive constant term, which would be unexplained by the simple additive model of Section 3 in which an additional unit of supply anywhere in the world has the same impact on the CME price. Market shares are nonnegative, and statistically significant rain impact coefficients in Tables XI and XII are negative. Therefore, a positive constant term would indicate that a region should achieve a certain minimum market share before having an impact on the CME market. The size of this constant term is about a tenth of the magnitude of the slope coefficient that can be interpreted as the rain impact for a region with 100% market share. It is in this sense that the estimated linear relationship is close to the prediction of the model in Section 3.

Two final caveats apply. If rain precipitation in the soybean producing regions in Brazil, which is not considered explicitly in this paper due to the lack of good quality data and geographical distance to Argentina, is in reality positively correlated with rain in Argentina, then the interpretation of the results outlined above should be modified accordingly. In this case, the effective market share that should be associated with rain

in Argentina should include some fraction of the output grown in Brazil. Therefore, the estimate of the slope in the regression of Table XIII would likely become even more negative, and the estimate of the constant term even more positive. Finally, corn, an imperfect substitute for soybeans, is also sensitive to rain and grown in the US Midwest and Argentina. Corn and soybean prices are partially linked through an economic relationship and therefore rain in any of these two regions could have an impact on soybean prices through the indirect channel of corn supply. This would not affect the results in Tables VIII to XII. The interpretation of rain impacts in terms of market shares, however, would require a multi-commodity model to take into account that the impact of rain in a specific region would depend on local market shares of corn and soybeans.

7. Conclusions

I study in this paper how exogenous local supply shocks in the distributed production of soybeans were incorporated into CME prices. Shocks were caused by rain precipitation in the central region of Argentina and the US Midwest, which combined account for close to 45% of global production. I generally find that rain precipitation had an economically important negative price impact, consistent with the fact that rain increases output and this increase in supply should lead to a decrease in prices. I also find that the price impact of rain, across regions and periods, is close to linearly related with the share of global output produced locally. This is generally consistent with the prediction of a very simple model that postulates a globally integrated market for soybeans in which an additional unit of supply anywhere in the world has the same impact in the CME price. Risk managers and speculators, in the US and elsewhere, are therefore exposed to globally distributed supply shocks.

References

- Bickel, J.E., Floehr, E., & Kim, S.D., (2011). Comparing NWS PoP forecasts to third-party providers. *Monthly Weather Review*, 139, 3304-3321
- Boudoukh, J., Richardson, M.P., Shen, Y.Q., & Whitelaw, R., (2007). Do asset prices reflect fundamentals? freshly squeezed evidence from the FCOJ market. *Journal of Financial Economics*, 83, 397-412.
- Chincarini, L., (2011). No chills or burns from temperature surprises: An empirical analysis of the weather derivatives market. *Journal of Futures Markets*, 31, 1-33.
- Chng, M.T., (2009). Economic linkages across commodity futures: hedging and trading implications. *Journal of Banking & Finance*, 33, 958-970.
- Christofoletti, M. A., Silva, R., & Mattos, F. (2012). The increasing participation of China in the world soybean market and its impact on price linkages in futures markets. Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St.Louis, MO. www.farmlandoc.illinois.edu/nccc134
- Deodhar, S., & Sheldon, I.M., (1997). Market power in the world market for soybean exports. *Journal of Agricultural and Resource Economics* 22, 78-86.
- Findlay, R., & O'Rourke, K.H., (2003). Commodity market integration. In: Bordo, M.D., Taylor, A.M., Williamson, J.G. (Eds.), *Globalization in Historical Perspective*, University of Chicago Press. <http://www.nber.org/chapters/c9585>.
- Fleming, J., Kirby, C., & Ostdiek, B., (2006). Information, trading, and volatility: Evidence from weather-sensitive markets. *The Journal of Finance* 61, 2899-2930.
- Frankel, J.A., & Rose, A.K., (2009). Determinants of agricultural and mineral commodity prices. Working paper series rwp10-038, Harvard University, Kennedy School of Government.
- Fung, H.G., Liu, Q., & Tse, Y., (2010). The information flow and market efficiency between the U.S. and chinese aluminum and copper futures markets. *Journal of Futures Markets* 30, 1192-1209.
- Hutson, S.S., Barber, N.L., Kenny, J.F., Linsey, K.S., Lumia, D.S., & Maupin, M.A., (2004). Estimated use of water in the United States in 2000. *US Geological Survey Circular* 1268.
- Isengildina-Massa, O., Irwin, S.H., Good, D.L., & Gomez J.K., (2008). The impact of situation and outlook information in corn and soybean futures markets: Evidence from WASDE reports. *Journal of Agricultural and Applied Economics* 40, 89-103.
- Regunaga, M. 2010. The soybean chain in Argentina. Implications of the organization of the commodity production and processing industry case studies.

Washington DC: World Bank, Latin America and the Caribbean Chief Economist Office.

Roberts, M.J., & Schlenker, W., (2010). Identifying supply and demand elasticities of agricultural commodities: Implications for the US ethanol mandate. NBER Working Paper 16385. <http://www.nber.org/papers/w15921>.

Roll, R., (1984). Orange juice and weather. *American Economic Review* 74, 861-880.

Tang, K., & Xiong, W., (2012). Index investment and financialization of commodities. *Financial Analysts Journal* 68, 54-74.

Tannura, M.A., Irwin, S.H., & Good, D.L., (2008). Weather, technology, and corn and soybean yields in the U.S. corn belt. Marketing and Outlook Research Report 2008-01. Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign.

Yang, J., Bessler, D. A., & Leatham, D. J. (2000). The law of one price: developed and developing country market integration. *Journal of Agricultural and Applied Economics* 32, 429-440.

	World Prod.	US Prod.	US Share	US Yield	Arg. Prod.	Arg. Share	Arg. Yield	Brl. Prod.	Brl. Share
96/97	132.30	64.78	0.49	2.53	11.20	0.08	1.81	27.30	0.21
97/98	158.24	73.18	0.46	2.62	19.50	0.12	2.80	32.50	0.21
98/99	159.83	74.60	0.47	2.62	20.00	0.13	2.45	31.30	0.20
99/00	160.35	72.22	0.45	2.46	21.20	0.13	2.47	34.70	0.22
00/01	175.76	75.06	0.43	2.56	27.80	0.16	2.67	39.50	0.22
01/02	184.82	78.67	0.43	2.66	30.00	0.16	2.63	43.50	0.24
02/03	196.87	75.01	0.38	2.56	35.50	0.18	2.82	52.00	0.26
03/04	186.64	66.78	0.36	2.28	33.00	0.18	2.36	51.00	0.27
04/05	215.78	85.02	0.39	2.84	39.00	0.18	2.71	53.00	0.25
05/06	220.67	83.51	0.38	2.90	40.50	0.18	2.66	57.00	0.26
06/07	236.23	87.00	0.37	2.88	48.80	0.21	2.99	59.00	0.25
07/08	220.47	72.86	0.33	2.81	46.20	0.21	2.82	61.00	0.28
08/09	211.96	80.75	0.38	2.67	32.00	0.15	2.00	57.80	0.27
09/10	260.84	91.42	0.35	2.96	54.50	0.21	2.93	69.00	0.26
10/11	261.97	90.61	0.35	2.92	49.50	0.19	2.66	67.50	0.26
Mean	198.85	78.10	0.40	2.68	33.91	0.16	2.59	49.07	0.24

Table I: Soybean annual production, in millions of metric tons, for the World, US, Argentina and Brazil, corresponding global market shares, and yields (in tons per hectare). Data was obtained from the USDA *Oilseeds Report* for December 2006, December 2010, and May 2011. 96/97 refers to the October 1996 harvest for the US, and the April 1997 harvest for the Southern hemisphere.

	Bean exports			Meal exports		
	World	US	Argentina	World	US	Argentina
96/97	36.76	24.11	0.72	30.25	6.45	8.35
97/98	39.63	23.76	3.21	33.26	8.72	12.98
98/99	37.93	21.90	3.10	35.38	6.98	12.76
99/00	45.63	26.54	4.11	34.19	6.91	12.85
00/01	53.66	27.10	7.38	36.26	7.34	15.98
01/02	52.90	28.95	6.22	41.81	7.27	17.57
02/03	61.24	28.42	8.81	43.07	5.73	19.16
03/04	56.04	24.13	6.80	46.09	4.69	19.08
04/05	64.82	29.86	10.69	47.70	6.66	22.70
05/06	63.43	25.58	7.13	52.52	7.30	24.72
06/07	70.86	30.39	12.13	55.32	7.99	28.11
07/08	78.78	31.54	11.80	56.60	8.38	24.39
08/09	76.84	34.82	3.49	52.80	7.71	21.30
09/10	92.65	40.85	13.67	55.72	10.14	27.82
10/11	95.62	42.18	9.50	59.64	8.30	29.25
Mean	61.79	29.34	7.25	45.37	7.37	19.80

Table II: Soybean and soymeal annual exports, in millions of metric tons, for the World, US, and Argentina. Data was obtained from the USDA *Oilseeds Report* for December 2006, December 2010, and May 2011. 96/97 refers to the October 1996 harvest for the US, and the April 1997 harvest for the Southern hemisphere.

Year	Bs. As.	Cdba.	E. Rios	Sta. Fe	Argentina	Share Central	Yield Central
96/97	2.53	2.91	0.28	4.16	11.00	0.90	1.69
97/98	3.86	5.82	0.73	7.31	18.73	0.95	2.76
98/99	4.58	5.26	0.76	7.30	19.96	0.90	2.49
99/00	3.78	6.93	0.54	6.64	20.12	0.89	2.36
00/01	5.73	8.15	1.66	8.66	26.87	0.90	2.68
01/02	5.78	9.66	1.91	8.35	29.96	0.86	2.70
02/03	7.14	9.85	2.81	10.22	34.78	0.86	2.92
03/04	7.85	8.38	2.31	9.14	31.51	0.88	2.34
04/05	10.00	11.19	3.05	10.45	38.22	0.91	2.93
05/06	10.53	11.12	2.80	10.28	40.47	0.86	2.74
06/03	11.65	14.17	3.93	11.30	47.38	0.87	3.10
07/08	12.25	12.75	3.29	11.48	46.16	0.86	2.95
08/09	6.74	11.17	1.14	8.08	30.94	0.88	1.99
09/10	17.05	12.99	4.03	10.43	52.61	0.85	2.96
Mean	7.82	9.31	2.09	8.14	32.05	0.88	2.62

Table III: Soybean annual production, in millions of metric tons, for the four provinces in the central region of Argentina. Also reported: total national output, share of national output produced by the central region of Argentina, and corresponding yield (tons per hectare). Data was obtained from the Agricultural Information Data Center of Argentina's Ministry of Agriculture, Livestock and Fisheries. <http://www.siaa.gov.ar/index.php/series-por-tema/agricultura>.

Year	IL	IN	IA	MN	MO	NE	OH	US	Share Midwest	Yield Midwest
96/97	10.86	5.54	11.32	6.10	4.08	3.69	4.28	64.84	0.71	2.69
97/98	11.65	6.28	13.02	6.95	4.75	3.91	5.20	73.18	0.71	2.85
98/99	12.63	6.29	13.52	7.77	4.63	4.49	5.26	74.61	0.73	2.91
99/00	12.06	5.89	13.02	7.70	4.00	4.92	4.41	71.94	0.72	2.69
00/01	12.51	6.86	12.65	7.98	4.76	4.73	5.08	75.06	0.73	2.82
01/02	13.01	7.46	13.08	7.25	5.07	6.07	5.11	78.68	0.72	2.89
02/03	12.24	6.42	13.47	8.41	4.63	4.80	3.85	74.30	0.72	2.78
03/04	10.33	5.55	9.33	6.49	3.97	4.96	4.49	66.78	0.68	2.35
04/05	13.47	7.74	13.54	6.33	6.08	5.95	5.65	85.02	0.69	3.11
05/06	12.09	7.18	14.50	8.33	5.00	6.41	5.49	84.01	0.70	3.18
06/07	13.13	7.73	13.88	8.78	5.29	6.82	5.91	87.01	0.71	3.18
07/08	9.80	6.00	12.21	7.28	4.77	5.34	5.42	72.87	0.70	3.08
08/09	11.64	6.65	12.11	7.19	5.20	6.15	4.39	80.54	0.66	2.90
09/10	11.71	7.26	13.23	7.75	6.28	7.06	6.04	91.43	0.65	3.18
10/11	12.69	7.04	13.51	8.95	5.73	7.29	6.00	90.62	0.68	3.27
Mean	11.99	6.66	12.83	7.55	4.95	5.51	5.10	78.06	0.70	2.93

Table IV: Soybean annual production, in millions of metric tons, for states in the US Midwest, defined as the combined output of IL, IN, IA, MN, MO, NE and OH. Also reported: total national output, share of national output produced by the US Midwest, and corresponding yield (tons per hectare). Data is taken from the USDA *Agricultural Statistics Report* for 1997, 2000, 2003, 2006, 2009, and 2011.

Total midsummer rain				
	Avg 97-03	Std dev 97-03	Avg 04-10	Std dev 04-10
Aeroparque	6.68	4.81	9.76	4.39
Ceres	10.07	7.01	7.96	3.20
Cordoba	9.42	3.66	9.03	3.88
Junin	8.18	4.12	10.59	4.76
Parana	7.79	3.72	9.75	5.62
Rio Cuarto	6.61	4.73	15.56	6.75
Rosario	7.84	3.20	10.27	7.50
Tandil	7.21	4.21	7.55	6.94

Average midsummer temperature				
	Avg 97-03	Std dev 97-03	Avg 04-10	Std dev 97-03
Aeroparque	74.7	2.1	75.5	0.7
Ceres	75.9	1.6	77.8	1.1
Cordoba	72.4	1.7	73.2	1.2
Junin	71.5	2.5	72.4	0.6
Parana	75.3	1.9	76.4	1.2
Rio Cuarto	72.3	2.0	72.3	1.0
Rosario	75.5	2.2	75.3	1.3
Tandil	67.7	2.1	68.0	2.2

Table V: Summary statistics for total midsummer rain and average midsummer temperature at weather stations in central Argentina. Rain measured in inches, temperature in degrees Fahrenheit. Data from the National Climatic Data Center administered by the National Oceanic and Atmosphere Administration.

Total midsummer rain				
	Avg 96-02	Std dev 96-02	Avg 03-10	Std dev 03-10
Greater Peoria IL	6.00	1.43	6.60	1.63
Springfield IL	6.59	1.51	6.55	2.35
Scott AFB IL	7.35	2.21	7.62	3.45
Des Moines IW	6.84	2.23	9.00	3.70
Waterloo IW	8.32	4.39	9.05	4.21
Sioux City IW	6.91	3.35	7.59	2.65
Omaha NE	8.27	4.66	8.15	3.05
Lincoln NE	6.85	2.38	5.72	1.88
Toledo OH	4.66	1.93	8.10	2.41
Columbus OH	7.03	1.63	8.67	2.30
Kansas MO	5.17	2.57	8.25	3.72
Whiteman MO	7.92	4.75	8.48	5.20
St Cloud MN	6.70	2.34	6.29	1.87
Rochester MN	9.21	3.01	8.34	3.53
Purdue IN	5.69	2.25	7.32	2.00
Grissom IN	7.34	2.86	9.57	4.83
South Bend IN	6.46	1.61	8.79	4.07
Average midsummer temperature				
	Avg 96-02	Std dev 96-02	Avg 03-10	Std dev 03-10
Greater Peoria IL	74.7	1.5	74.4	2.7
Springfield IL	74.8	1.1	74.5	2.7
Scott AFB IL	77.6	1.1	76.6	2.5
Des Moines IW	74.7	1.4	75.0	2.7
Waterloo IW	72.3	1.2	71.7	2.4
Sioux City IW	73.3	1.4	73.8	2.3
Omaha NE	75.9	1.5	75.8	2.7
Lincoln NE	76.7	1.8	76.6	2.3
Toledo OH	72.2	2.2	72.3	1.9
Columbus OH	74.4	1.9	74.4	1.9
Kansas MO	77.6	1.8	77.8	3.0
Whiteman MO	79.7	2.5	76.3	2.9
St Cloud MN	69.4	1.6	69.7	2.6
Rochester MN	69.1	1.5	69.5	2.6
Purdue IN	73.8	1.2	73.3	2.2
Grissom IN	73.3	3.7	74.0	2.5
South Bend IN	72.1	2.1	71.9	2.2

Table VI: Summary statistics for total midsummer rain and average midsummer temperature at weather stations in the US Midwest. Rain measured in inches, temperature in degrees Fahrenheit. Data from the National Climatic Data Center administered by the National Oceanic and Atmosphere Administration.

Estimation results by OLS regressions of $Y_i^{Region} = \alpha^{Region} R_i^{Region} + Const + \epsilon_i$. Yields, midsummer and extended summer rain, with annual frequency, 1996/7-2010. By *, ** and *** I indicate statistical significance at the 10%, 5% and 1% levels.

	Annual yield US Midwest	Annual yield Central Argentina
Midsummer rain	0.161*** (0.048)	0.117*** (0.043)
Const	1.730*** (0.357)	1.508*** (0.421)
R^2	0.467	0.402
	Annual yield US Midwest	Annual yield Central Argentina
Extended summer rain	0.030* (0.020)	0.042** (0.017)
Const	2.268*** (0.445)	1.447*** (0.492)
R^2	0.146	0.350

Table VII: Impact of midsummer (top panel) and extended summer (bottom panel) rain on regional yield.

Estimation results by OLS regression with Newey-West standard errors of

$$\frac{P_i^{May} - P_{i-1}^{May}}{P_{i-1}^{May}} = \alpha + \beta_{Arg} Rain_i^{Arg} + controls_i + \epsilon_i,$$

for the effect of summer rain in Argentina on the daily return of the CME May Soybean future price. The sample for a fixed May contract includes rain in the previous January and February. Sample: Monday to Friday, from the first day in January to the last day in February, between 1997 to 2010. I also report results from the estimation of

$$\frac{P_i^{Nov} - P_{i-1}^{Nov}}{P_{i-1}^{Nov}} = \alpha + \beta_{US} Rain_i^{US} + controls_i + \epsilon_i,$$

for the effect of summer rain in the US on the daily return of the CME November Soybean future price. The sample for a fixed November contract includes rain in the previous July and August. Sample: Monday to Friday, from the first day in July to the last day in August, between 1996 to 2010. The variables $Rain_i^{Arg}$, $Rain_i^{US}$ and the Temperature controls are spatial averages in the region of interest over four consecutive days from day $i + 1$. Rain data includes weekends. All other controls are returns between day $i - 1$ and day i . By *, ** and *** I indicate statistical significance at the 10%, 5% and 1% levels.

	May contracts 97 - 03	May contracts 04 - 10	Nov. contracts 96 - 02	Nov. contracts 03 - 10
β_{Arg}	0.0029 (0.0038)	-0.0153*** (0.0045)		
β_{US}			-0.0524*** (0.0113)	-0.0184* (0.0115)
Oil	-0.0021 (0.0251)	0.1622*** (0.0400)	-0.0355 (0.0436)	0.2615*** (0.0550)
FXRate	0.2067 (0.2271)	-0.6388** (0.2852)	-0.2926 (0.2818)	-0.3975 (0.2993)
BDI	-0.0125 (0.0547)	0.0725** (0.0320)	-0.0307 (0.1914)	-0.0622 (0.0507)
Temp Arg	-0.0002 (0.0001)	-0.0003 (0.0003)		
Temp US			0.0000 (0.0003)	0.0002 (0.0002)
Hog	0.0147 (0.0389)	0.0614 (0.0752)	0.1038 (0.0645)	-0.0983 (0.0825)
Const	-0.0007 (0.0009)	0.0037*** (0.0013)	0.0061*** (0.0016)	0.0018 (0.0015)
Observations	297	295	311	353
R-squared	0.0105	0.1462	0.0613	0.1010

Table VIII: Impact of midsummer rain in Argentina and US Midwest on CME Soybean future prices.

Estimation results by OLS regression with Newey-West standard errors of

$$\frac{P_i - P_{i-1}}{P_{i-1}} = \alpha + \beta_{Arg} Rain_i^{Arg} + \beta_{US} Rain_i^{US} + controls_i + \epsilon_i,$$

for the joint effect of rain precipitation in Argentina and the US on the daily return of CME Soybean future prices. Price changes are computed for the May and November contracts. The sample for any fixed contract includes rain in the previous January, February, July and August (possibly from the previous year), corresponding to both summer and winter rain in Argentina and the US. The variables $Rain_i^{Arg}$, $Rain_i^{US}$ and the Temperature controls are spatial averages in the region of interest over four consecutive days from day $i + 1$. Sample for May contract returns: Monday to Friday, in January, February, July and August, from July 1996 to February 2010. Sample for the November contract returns: Monday to Friday, in January, February, July and August, from January 1997 to August 2010. Rain data includes weekends. All other controls are returns between day $i - 1$ and day i . By *, ** and *** I indicate statistical significance at the 10%, 5% and 1% levels.

	May contracts 97 - 03	May contracts 04 - 10	Nov. contracts 97 - 02	Nov. contracts 03 - 10
β_{Arg}	-0.0048 (0.0034)	-0.0092** (0.0041)	-0.0046 (0.0040)	-0.0065* (0.0037)
β_{US}	-0.0295*** (0.0078)	-0.0161* (0.0090)	-0.0328*** (0.0090)	-0.0159** (0.0078)
Oil	-0.0171 (0.0218)	0.1906*** (0.0328)	-0.0297 (0.0250)	0.1861*** (0.0303)
FXRate	-0.0250 (0.1669)	-0.6033*** (0.2204)	-0.1261 (0.1898)	-0.4949** (0.1974)
BDI	-0.0214 (0.0619)	0.0352 (0.0314)	0.0067 (0.0707)	-0.0092 (0.0329)
Temp Arg	0.0000 (0.0001)	0.0001 (0.0001)	0.0000 (0.0001)	0.0000 (0.0001)
Temp US	-0.0001 (0.0001)	0.0002** (0.0001)	-0.0001 (0.0001)	0.0002** (0.0001)
Hog	0.0406 (0.0358)	0.0215 (0.0550)	0.0680 (0.0413)	-0.0506 (0.0497)
Const	0.0029*** (0.0010)	0.0027** (0.0011)	0.0032*** (0.0011)	0.0025** (0.001)
Observations	608	604	520	691
R-squared	0.0355	0.1106	0.0363	0.0962

Table IX: Joint impact of rain in Argentina and US Midwest on CME Soybean future prices.

Estimation results by OLS regression with Newey-West standard errors of

$$\frac{P_i^{May} - P_{i-1}^{May}}{P_{i-1}^{May}} = \alpha + \beta_{Arg} Rain_i^{Arg} + controls_i + \epsilon_i,$$

for the effect of summer rain in Argentina on the daily return of the CME May Soybean future price. The sample for a fixed May contract includes rain in the previous November-April period. Sample: Monday to Friday, from the first day in November, to the last day in April of the following year, for 1996 to 2009. I also report results from the estimation of

$$\frac{P_i^{Nov} - P_{i-1}^{Nov}}{P_{i-1}^{Nov}} = \alpha + \beta_{US} Rain_i^{US} + controls_i + \epsilon_i,$$

for the effect of summer rain in the US on the daily return of the CME November Soybean future price. The sample for a fixed November contract includes rain in the previous May-October period. Sample: Monday to Friday, from the first day in May, to the last day in October, for 1996 to 2010. The variables $Rain_i^{Arg}$, $Rain_i^{US}$ and the Temperature controls are spatial averages in the region of interest over four consecutive days from day $i + 1$. Rain data includes weekends. All other controls are returns between day $i - 1$ and day i . By *, ** and *** I indicate statistical significance at the 10%, 5% and 1% levels.

	May contracts 97 - 03	May contracts 04 - 10	Nov. contracts 96 - 02	Nov. contracts 03 - 10
β_{Arg}	-0.0012 (0.0022)	-0.0064** (0.0031)		
β_{US}			-0.0186*** (0.0060)	-0.0093** (0.0046)
Oil	0.0170 (0.0110)	0.1623*** (0.0224)	-0.0009 (0.0183)	0.2772*** (0.0262)
FXRate	0.0482 (0.1203)	-0.8089*** (0.1564)	-0.1676 (0.1384)	-0.6492*** (0.1921)
BDI	0.0149 (0.0394)	0.0171 (0.0290)	-0.0573 (0.0715)	-0.0056 (0.0266)
Temp Arg	-0.0001 (0.0001)	-0.0001 (0.0001)		
Temp US			0.0000 (0.0001)	0.0000 (0.0001)
Hog	0.0285 (0.0216)	0.0678* (0.0390)	0.0628** (0.0295)	-0.0221 (0.0419)
Const	0.0003 (0.0005)	0.0019*** (0.0007)	0.0019** (0.0009)	0.0012 (0.0008)
Observations	905	907	921	1050
R-squared	0.0059	0.1446	0.0196	0.1665

Table X: Impact of extended summer rain in Argentina and US Midwest on CME Soybean future prices.

Estimation results by OLS regression with Newey-West standard errors of

$$\frac{P_i^{May} - P_{i-1}^{May}}{P_{i-1}^{May}} = \alpha + \beta_{Arg} Rain_i^{Arg} + \beta_{US} Rain_i^{US} + controls_i + \epsilon_i,$$

for the joint effect of rain precipitation in Argentina and the US on the daily return of CME May Soybean future prices. The sample for any fixed contract includes rain in the previous January, February, July and August (possibly from the previous year), corresponding to both summer and winter rain in Argentina and the US. The variables $Rain_i^{Arg}$, $Rain_i^{US}$ and the Temperature controls are spatial averages in the region of interest over four consecutive days from day $i + 1$. Sample for May contract returns: Monday to Friday, in January, February, July and August, from July 1996 to February 2010. Sample for the November contract returns: Monday to Friday, in January, February, July and August, from January 1997 to August 2010. Rain data includes weekends. All other controls are returns between day $i - 1$ and day i . By *, ** and *** I indicate statistical significance at the 10%, 5% and 1% levels. In the interest of space I only report estimates of the impact of rain.

Region and Period	β_{Arg}	β_{US}	N obs.	R-squared
97-98	0.0021 (0.0066)	-0.0358*** (0.0152)	174	0.08
99-00	0.0057 (0.0125)	-0.0269 (0.0165)	171	0.06
01-02	-0.0099** (0.0049)	-0.0366*** (0.0110)	175	0.10
03-04	-0.0034 (0.0072)	-0.0164 (0.0129)	174	0.06
05-06	-0.0179*** (0.0063)	-0.0324** (0.0157)	171	0.09
07-08	0.0026 (0.0091)	0.0001 (0.0177)	176	0.16
09-10	-0.0140* (0.0076)	-0.0284 (0.0221)	171	0.16

Table XI: Impact of midsummer rain in Argentina and US Midwest on CME May Soybean future price, over biennial periods.

Estimation results by OLS regression with Newey-West standard errors of

$$\frac{P_i^{Nov} - P_{i-1}^{Nov}}{P_{i-1}^{Nov}} = \alpha + \beta_{Arg} Rain_i^{Arg} + \beta_{US} Rain_i^{US} + controls_i + \epsilon_i,$$

for the joint effect of rain precipitation in Argentina and the US on the daily return of CME May Soybean future prices. The sample for any fixed contract includes rain in the previous January, February, July and August (possibly from the previous year), corresponding to both summer and winter rain in Argentina and the US. The variables $Rain_i^{Arg}$, $Rain_i^{US}$ and the Temperature controls are spatial averages in the region of interest over four consecutive days from day $i + 1$. Sample for May contract returns: Monday to Friday, in January, February, July and August, from July 1996 to February 2010. Sample for the November contract returns: Monday to Friday, in January, February, July and August, from January 1997 to August 2010. Rain data includes weekends. All other controls are returns between day $i - 1$ and day i . By *, ** and *** I indicate statistical significance at the 10%, 5% and 1% levels. In the interest of space I only report estimates of the impact of rain.

Region and Period	β_{Arg}	β_{US}	N obs.	R-squared
97-98	0.0044 (0.0067)	-0.0267* (0.0137)	173	0.08
99-00	0.0036 (0.0127)	-0.0313 (0.0223)	172	0.05
01-02	-0.0126** (0.0051)	-0.0194* (0.0112)	175	0.06
03-04	-0.0033 (0.0073)	-0.0239* (0.0144)	173	0.05
05-06	-0.0153** (0.0063)	-0.0274** (0.0115)	171	0.11
07-08	0.0024 (0.0090)	0.0008 (0.0218)	176	0.24
09-10	-0.0110* (0.0063)	-0.0113 (0.0149)	171	0.13

Table XII: Impact of midsummer rain in Argentina and US Midwest on CME November Soybean future price, over biennial periods.

Estimation results by OLS regression with Newey-West standard errors of

$$\beta_{Region,Period} = \delta S_{Region,Period} + Const + \epsilon_{Region,Period}.$$

The dependent variable is the impact of rain, defined as the OLS coefficient of the CME soybean price change on daily rain, in a certain period and region, or its renormalized form. Values are taken from Tables XI and XII, and reproduced below. The explanatory variable is the average share of global production for certain region and period, computed using data from Tables I, III and IV. The upper part of the table below presents the data, and the bottom part presents regression results. By *, ** and *** I indicate statistical significance at the 10%, 5% and 1% levels.

Region, period	Share	β CME May	β CME May. renorm.	β CME Nov.	β CME Nov. renorm.
Argentina 97-98	0.093	0.0021	0.0470	0.0044	0.0985
Argentina 99-00	0.116	0.0057	0.1276	0.0036	0.0806
Argentina 01-02	0.141	-0.0099	-0.2217	-0.0126	-0.2822
Argentina 03-04	0.157	-0.0034	-0.0761	-0.0033	-0.0739
Argentina 05-06	0.159	-0.0179	-0.4008	-0.0153	-0.3426
Argentina 07-08	0.182	0.0026	0.0582	-0.0024	-0.0537
Argentina 09-10	0.155	-0.0140	-0.3135	-0.0110	-0.2463
Midwest 97-98	0.335	-0.0358	-0.6515	-0.0267	-0.4859
Midwest 99-00	0.319	-0.0269	-0.4895	-0.0313	-0.5696
Midwest 01-02	0.292	-0.0366	-0.6661	-0.0194	-0.3531
Midwest 03-04	0.257	-0.0164	-0.2985	-0.0239	-0.4350
Midwest 05-06	0.264	-0.0324	-0.5896	-0.0274	-0.4986
Midwest 07-08	0.241	0.0001	0.0018	0.0008	0.0146
Midwest 09-10	0.233	-0.0284	-0.5168	-0.0113	-0.2056
Share of global prod.		-0.1529*** (0.0351)	-2.690*** (0.697)	-0.1226*** (0.0276)	-2.141*** (0.558)
Const		0.0171** (0.0078)	0.281*** (0.155)	0.0132** (0.0062)	0.211* (0.124)
Observations		14	14	14	14
R^2		0.61	0.55	0.62	0.55

Table XIII: Market shares and price impact of local rain.

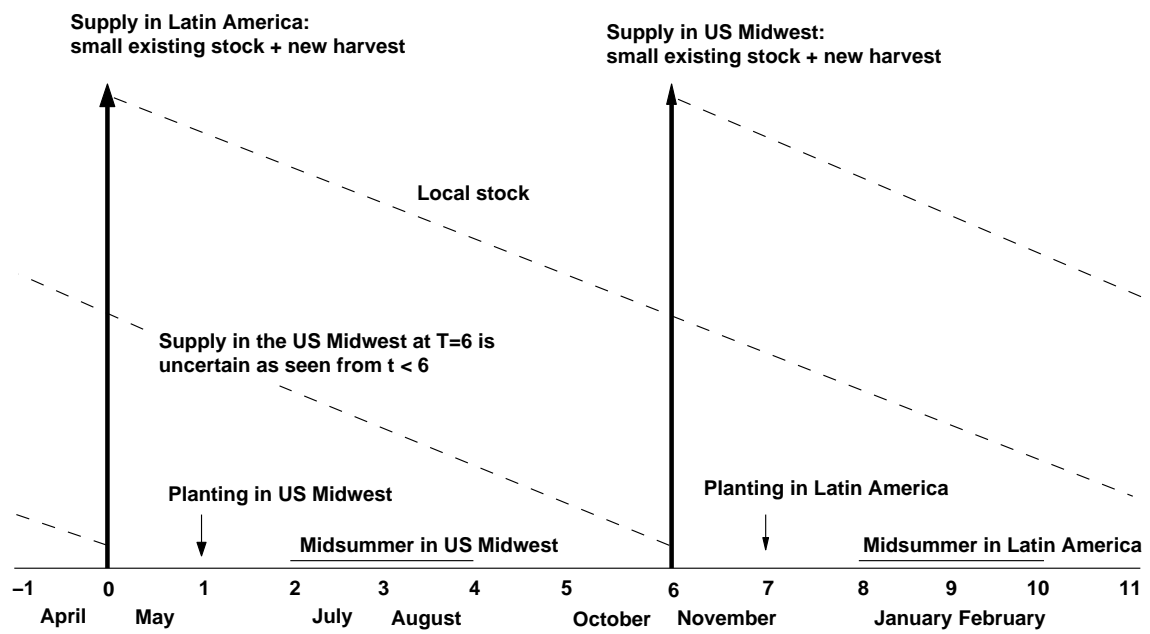


Figure 1: Time-line of events in the soybean global market.

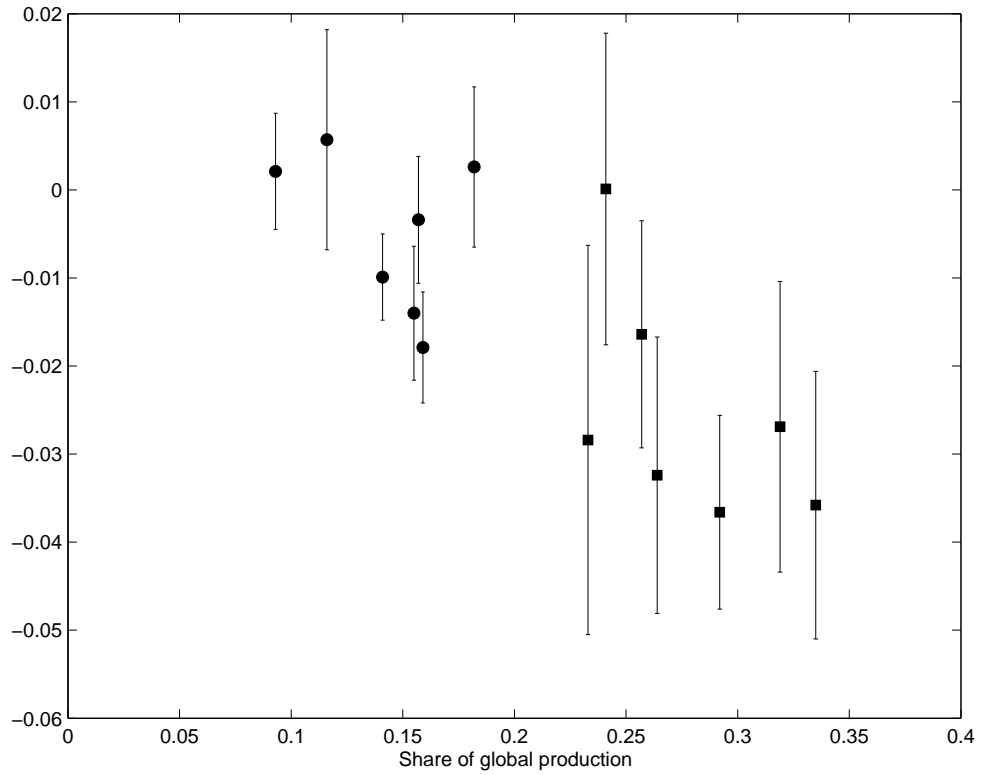


Figure 2: Impact of local rain on the CME May soybean price, as a function of share of global production. Circles and squares represent estimates for the central region of Argentina and the US Midwest respectively. Each point corresponds to a biennial period. Values are taken from Table XIII.

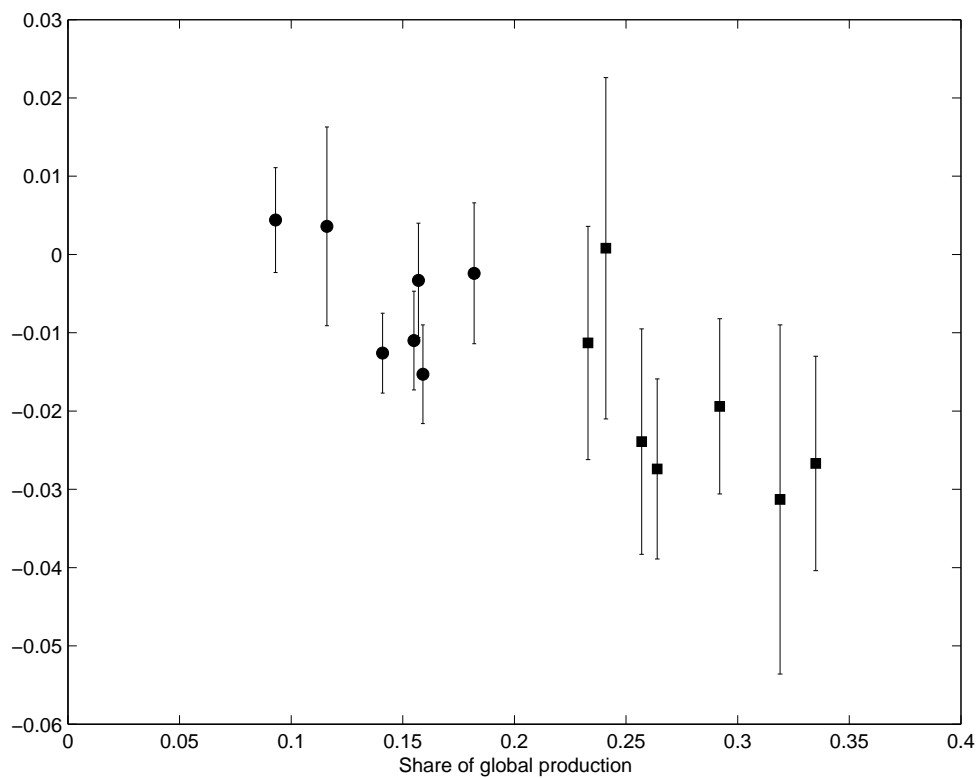


Figure 3: Impact of local rain on the CME November soybean price, as a function of share of global production. Circles and squares represent estimates for the central region of Argentina and the US Midwest respectively. Each point corresponds to a biennial period. Values are taken from Table XIII.