The Value of Public Information in Storable Commodity Markets: Application to the Soybean Market*

Christophe Gouel†

August 29, 2017

Abstract

Public information is routinely provided in commodity markets through reports on stock levels, demand levels, or production forecasts. The purpose of this study is to provide a framework to estimate the potential benefits of these data. This framework is applied to the case of the soybean market. A rational expectations storage model of the global soybean market accounting for both inter-annual and intra-annual market dynamics is built. Between planting and harvesting, shocks affect the size of the potential harvest. Estimates of the size of these shocks are publicly reported and affect the market equilibrium through adjustments in stock levels. The effects of the news shocks are analyzed by varying counterfactually the observability of the seasonal shocks. The presence of advance information has a limited effect on the inter-annual price volatility but redistributes price volatility during the season, increasing it just before the harvest when almost all news has been received but stocks are tight, and decreasing it after. The effect of news shocks is stronger on higher-order moments of the distribution with a strong decrease in skewness and kurtosis related to the lower frequency of price spikes. Inter-annual stock levels are also significantly decreased by the availability of advance information and their lower levels explain most of the welfare gains.

Keywords: Commodity price dynamics, commodity storage, news shocks.

JEL classification: D84, Q02, Q11, Q13.

*The author is grateful to Tamara Ben Ari, Joe Glauber, and Will Martin for helpful comments. This research was generously supported by the INRA-Cirad metaprogramme GloFoodS. Christopher Marciniak and Maïmouna Yokessa provided excellent research assistance.

†Économie Publique, AgroParisTech, INRA, Université Paris-Saclay, 78850, Thiverval-Grignon, France and IFPRI, 2033 K St NW, Washington, DC 20006, USA. Email: christophe.gouel@inra.fr

1 Introduction

Accurate and timely information is key to the good functioning of financial markets and commodity markets are no exception. What is specific to commodity markets is the extensive public involvement in the provision of market information. For example, national or international agencies regularly report the levels of stocks and, for agricultural commodities, forecasts of the coming harvests. Recently, improving the transparency of the global food market was the rationale for the creation of the Agricultural Market Information System (AMIS) in 2011 by the G20 Ministers of Agriculture following the global food price hikes in 2007/08 and 2010. The important role of public information is often acknowledged in the literature (C-FARE, 2016), however, evaluations of its effects and of the benefits of its existence are rare.

The objective of this article is to fill this gap by developing a framework to assess the economic role of public information in storable commodity markets. Public information can affect market allocation and welfare in many ways (see C-FARE, 2016, for a recent survey): by allowing an efficient market allocation, by leveling the playing field between market participants with unequal access to private information, or by constituting the basis for research on market behavior. Here, we focus on the reduction in misallocation of resources in time. For this, we build a rational expectations storage model that accounts for both inter-annual and intra-annual market dynamics. In the model, shocks between planting and harvesting affect the size of the potential harvest. We apply the model to the case of soybeans, because its production is very concentrated—more than 80% of world production comes from Argentina, Brazil, and the United States—and its market is well integrated internationally (for example, Merener, 2015, shows that soybean futures react similarly to weather events in Argentina, Brazil, or the United States), two features that simplify the modeling. To make the model tractable, we define a simplified crop calendar corresponding to the most representative planting and harvesting dates and we use monthly changes in U.S. Department of Agriculture (USDA) production forecasts to calibrate the size of the news shocks. By varying counterfactually the size of the seasonal shocks that are observable, we can assess what are the benefits of providing public information about them.

Because of its focus on misallocation of resources in time, one can think of the present article as a modernization of the approaches proposed in Hayami and Peterson (1972) and Bradford and Kelejian (1977, 1978). These papers developed analytical frameworks based on intra-annual storage models to assess the welfare costs of measurement errors in public information. A key difference with our approach is that we include the problem of seasonal news shocks into a modern rational expectations storage model, so in our case the seasonal information does not matter only for seasonal market dynamics, but also for inter-annual market dynamics, because advance information about the harvests allows smoother inter-annual stock adjustments. We have also a slightly different emphasis. These papers were interested in assessing the welfare benefits of reducing forecast errors. Because the forecast errors cannot be observed, we prefer to assume that the production forecasts are made without error, although we acknowledge that part of the production shocks can be unobserved. By using the last 40 years of data from USDA, we can quantify the size of the informational shock at each period. However, adopting the framework of the modern rational expectations model has drawbacks. This type of models is not analytically tractable, so we lose some of the insights provided by the analytical simplicity of Hayami and Peterson's framework. But despite its reliance on numerical simulations, we have aimed to keep our modeling framework as simple as possible so that all results have clear interpretations. For simplicity, we will neglect one potential source of welfare gains associated with public information and considered in Hayami and Peterson (1972): the adjustments of production to news shocks. We assume that supply is elastic at planting time, but after planting it is not possible for production to react to the shocks (see Bontems and Thomas, 2000, and Lechthaler and Vinogradova, 2017, for work on this issue).

Since Hayami and Peterson (1972) and Bradford and Kelejian (1977, 1978), the literature on this topic has mostly moved to another approach inspired by the market efficiency studies in finance and which consists in assessing if and how markets react to announcements (e.g., Garcia et al., 1997). This more recent literature addresses the question of the value of USDA forecasts through their capacity to affect the market. The USDA market outlooks are not the only source of market information. Large trading firms gather their own information about market fundamentals. Some private companies have specialized in selling market outlook in advance of USDA releases. The existence of other, private, sources of information raises questions about the economic role of public information. One way to assess its role is to quantify the market reaction to the release of the USDA reports. Adjemian (2012b) and Adjemian et al. (2016) show that agricultural markets do indeed react to the information contained in the USDA reports and that this reaction is consistent with storage theory.[1] While this literature can quantify how markets react to the release of public information, it is silent on the benefits of this information, which is the object of the present study. However, for the tractability of our rational expectations model, we cannot account for the existence of various sources of information. We assume that all agents have the same set of information. So, we do not test whether USDA reports move the market by providing additional information, we rather assume that USDA reports constitute a revelation of common information. That part of this information could have been incorporated a few weeks before because agents have had access to private information is irrelevant in our monthly framework where each month coincides with the revelation of new information from USDA.

We use USDA production forecasts because their monthly releases are closely watched by market observers and are considered as a benchmark. USDA started producing monthly crop reports in 1863 during the Civil War when timely information on crop conditions where crucial for supplying the army (Adjemian, 2012a). Since then, with very few exceptions, USDA released reports on crop production around the tenth of every month. The "Great Grain Robbery" of 1972, when the Soviet Union purchased large amounts of subsidized U.S. grains before world food prices spiked, led to the establishment of an interagency process establishing balance sheets, so consolidating together information on supply and demand. Here, we use only the production forecasts because, after planting, changes in production forecasts can reasonably be considered as exogenous and so as news shocks, while the other dimensions of the balance sheets (domestic use, exports, and ending stocks) mix endogenous market responses and news shocks. The most important foreign markets for each

[1] Karali and Thurman (2009) shows also that the lumber futures market reacts to announcements in a way consistent with the storage model.

commodity started to be added to the balance sheets in 1980, which provides us with changes in production forecasts for Argentina and Brazil. Now, other institutions produce crop forecasts, either private firms, or international organizations such as the International Grains Council or, since 2011, AMIS, but there is no evidence that their forecasts move markets in any significant way.

Our starting point is the rational expectations storage model extended to account at the same time for inter-annual and intra-annual market dynamics. Adopting the rational expectations framework is a natural starting point as it implies that agents optimally process the information provided by the news shocks.[2] Starting with the early work of Lowry et al. (1987), Williams and Wright (1991, Ch. 8), and Chambers and Bailey (1996), a few papers have built rational expectations model with intra-annual dynamics. Only three (Ng and Ruge-Murcia, 2000; Osborne, 2004; Peterson and Tomek, 2005) include news shocks, where the word "news" takes the sense it has been given in macroeconomic literature (Beaudry and Portier, 2014) of observation of exogenous shocks with effects on fundamentals in the future. Ng and Ruge-Murcia (2000) show how news shocks can be a useful extension of the model to increase the serial correlation in prices, a long-standing puzzle since the estimations by Deaton and Laroque (1992, 1996).[3] Peterson and Tomek (2005) introduce news shocks in a model of the U.S. corn market to improve its capacity to explain observed seasonal market behavior. Osborne (2004), which is the closest work to ours, studies how precipitations observed in advance of the harvest affect the market equilibrium in a storage model of the Ethiopian grain market. She shows how the observation of news affect market equilibrium compared to a situation without news. Compared to Osborne (2004), we develop a richer model with a more detailed structure of news shocks and that includes an elastic supply. In particular, by adopting a monthly model, instead of Osborne's quarterly model, we can show that news shocks affect price volatility in subtle ways: they do not change much the average inter-annual price volatility but redistribute intra-annual price volatility, increasing it just before harvest when the harvest is almost known and stocks are at their minimum and decreasing it significantly otherwise. In addition, because we rely on the USDA production forecasts, we do not have to build the news shocks from meteorological information as we have direct observation of the news shocks expressed in quantities, which allows a better identification of the information available to the agents. Lastly, a related problem of information frictions, but not news in the previous sense, has been recently addressed in a rational expectations storage model by Steinwender (forthcoming). She shows how the establishment of the transatlantic telegraph reduced the information frictions related to the international trade of cotton between Great Britain and the United States. Since transatlantic shipping takes time, before the telegraph, prices in the export and import markets were based on lagged information from the other market. She estimates that the instantaneous flow of information allowed by the telegraph generated efficiency gains equivalent to 9% of export value. In our case, the welfare gains arise mostly from a better intertemporal allocation of resources and amount to 2% of the value of storage cost.

[2]The recent work of Abbott et al. (2016) shares objectives similar to ours but without rational expectations, which makes unclear how agents process the new information.

[3]Although recent works (Cafiero et al., 2011; Gouel and Legrand, 2017) have come a long way towards solving this puzzle.

Finally, our work can be related to the macroeconomic literature, where news shocks have been introduced into DSGE models and viewed as providing a positive theory of business cycle fluctuations (Beaudry and Portier, 2014). This is also one motivation for introducing news in the storage model (Ng and Ruge-Murcia, 2000; Peterson and Tomek, 2005), but this is not the only angle adopted in this article, which is also concerned with the welfare benefits associated to this information. Besides, it is worth noting that there are at least two key differences in the analysis of news shocks between macroeconomic models and the storage model. First, in agriculture while the underlying shocks are likely exogenous, the observability of these shocks as news shocks is partly a function of public intervention. Second, most news shocks in agriculture have a clear physical interpretation (e.g., size of planted area, expected bad yields), many of them being observables at some cost while in macroeconomic models news shocks are rarely observed and are either recovered from financial variables or model-based structural decomposition.

The rest of this paper is organized as follows. Section 2 develops our seasonal storage model. Section 3 describes how the model is taken to the data. It shows how USDA crop forecasts can be used to recover the size of the news shocks and describes new econometric approaches to estimate the storage cost and the supply elasticity. The estimation of the supply elasticity is done with a method inspired from Roberts and Schlenker (2013) but refined to account for the flows of information coming from the asynchronous harvests in the two hemispheres. Section 4 then presents the results of our counterfactual simulations. Section 5 offers some concluding remarks.

2 A seasonal storage model

Let consider a seasonal rational expectations storage model with 12 seasons per year. Years are indexed by t and seasons by $i \in \{1, \dots, I\}$ with the convention that $\{I + 1, t\} = \{1, t + 1\}$ and $\{1 - 1, t\} = \{I, t - 1\}$. There are two producing regions, indexed $r \in \{\text{US}, \text{LAC}\}$, corresponding to the United States, and Argentina and Brazil combined. The two regions differ by their planting and harvesting dates and the size of their potential production. We adopt the convention that the year is the U.S. marketing year and starts with the U.S. harvest. We assume that trade is costless so that price, $P_{i,t}$, is the same in all regions and the localization of final demand or storers does not matter for the equilibrium.

Soybeans generally takes around five months to grow from sowing to harvest. But because of the variety of local climates, sowing to harvesting extends in the United States from mid-April to late December. The same pattern can be observed in Argentina and Brazil. To simplify the model, we assume that the United States production takes place over the five most active months from May/June to October/November and the Argentina and Brazil production takes place from October/November to March/April. Because of this simplified crop calendar, the crop cycles do not overlap between regions, when planting occurs in one country the harvest has been completed in the other. Since producers can observe the size of the harvest from the other Hemisphere before planting, they can adjust their production based on this information.

2.1 Consumers

Final demand is given by a demand function, which is a downward sloping deterministic function of current price and is identical for every season: $D\left(P_{i,t}\right)$.

2.2 Storers

There is a single representative speculative storer, which is risk neutral and acts competitively. Its activity is to transfer a commodity from one period to the next. Storing the non-negative quantity $S_{i,t}$ from period $\{i,t\}$ to period $\{i+1,t\}$ entails a purchasing cost, $P_{i,t}S_{i,t}$, and a storage cost, $k\bar{P}S_{i,t}$, with k the unit physical cost of storage expressed in proportion of the steady-state annual price $\bar{P}$. The benefits in period $\{i+1,t\}$ are the proceeds from the sale of previous stocks: $P_{i+1,t}S_{i,t}$. The storer follows a storage rule that maximizes its expected profit which, accounting for the non-negativity constraint on stocks, leads to the following non-arbitrage condition

$$\beta\,\mathrm{E}_{i,t}\,P_{i+1,t} - P_{i,t} - k\bar{P} \leq 0, = 0 \text{ if } S_{i,t} > 0, \tag{1}$$

where $\mathrm{E}_{i,t}$ is the expectation operator conditional on period $\{i,t\}$ information and β is the monthly discount factor and is assumed to be fixed.

2.3 Producers

Production is undertaken in each region by a representative competitive producer with decreasing return to scale who takes the planting decision before knowing the selling price and the yield. The producer in region r plants in season i_π^r of year t with the expectation of harvesting five months later the quantity $Q^r_{i_\pi^r,t}$. However, the production is affected by a multiplicative random shock $\epsilon^r_{i_\pi^r+5,t}$ that follows a distribution with unitary mean described later. We assume that after planting the producer cannot adjust its production level. The producer chooses the production level by solving the following maximization of expected profit:

$$\max_{Q^r_{i_\pi^r,t}} \beta^5\,\mathrm{E}_{i_\pi^r,t}\left(P_{i_\pi^r+5,t}\epsilon^r_{i_\pi^r+5,t}Q^r_{i_\pi^r,t}\right) - \Psi^r\left(Q^r_{i_\pi^r,t}\right), \tag{2}$$

where $\Psi^r\left(Q^r_{i_\pi^r,t}\right)$ is the cost of planning the production $Q^r_{i_\pi^r,t}$ and $\epsilon^r_{i_\pi^r+5,t}Q^r_{i_\pi^r,t}$ is the realized production level. Profit maximization gives the following intertemporal equation

$$\beta^5\,\mathrm{E}_{i_\pi^r,t}\left(P_{i_\pi^r+5,t}\epsilon^r_{i_\pi^r+5,t}\right) = \Psi^{r\prime}\left(Q^r_{i_\pi^r,t}\right), \tag{3}$$

which equalizes the marginal cost of production and the expected discounted marginal benefit of one unit of planned production.

2.4 News shocks

In contrast with standard annual storage models, we do not assume that the production shock is concentrated in one period. During the growing season, potential production is affected by a series of shocks. We assume that these production shocks are partly observable after their realization. So, they provide advance information about the size of the coming harvest. They are news shocks in the sense that they do not directly affect current quantities, but expectations about realizations of future quantities. Nonetheless, they will affect market equilibrium through adjustments to stock levels.

The literature on news shocks has adopted two modeling strategies (Beaudry and Portier, 2014): to assume either a noisy signal composed of the true shock plus a noise or a shock composed of two elements, but only one being observable. In the small theoretical literature on news shocks in commodity markets, Hayami and Peterson (1972) and Bradford and Kelejian (1977, 1978) have adopted the approach of a noisy signal and Ng and Ruge-Murcia (2000), Osborne (2004), and Peterson and Tomek (2005) have adopted the other approach. With the noisy-signal approach, it is possible to generate richer dynamics as agents may have to react to false information in case of large noise shock. But the downside of this approach is that it requires a larger state space, because one has to follow separately the true shock and the associated noise, and in the case of our storage model it would lead to a model more difficult to calibrate, because there is no information about the size of the potential noise. So, we adopt the other approach.

The total production shock $\epsilon^r_{i^r_\pi+5,t}$ is assumed to be the product of a succession of five lognormally-distributed shocks occurring after planting:

$$\epsilon^r_{i^r_\pi+5,t} = \exp\left(\sum_{i=1}^{5} \eta^r_{i^r_\pi+i,t}\right), \tag{4}$$

where the $\eta^r_{i,t}$ are i.i.d. and follow normal distributions with mean μ^r_i and standard deviations σ^r_i such that $\mu^r_i + \left(\sigma^r_i\right)^2/2 = 0$, which ensures that seasonal shocks have unitary mean. Under these assumptions, $\epsilon^r_{i^r_\pi+5,t}$ follow a lognormal distribution with parameters $\sum_{i=1}^{5} \mu^r_{i^r_\pi+i,t}$ and $\sqrt{\sum_{i=1}^{5}\left(\sigma^r_{i^r_\pi+i,t}\right)^2}$. Given that $\mu^r_i + \left(\sigma^r_i\right)^2/2 = 0$, it implies that the expected mean of $\epsilon^r_{i^r_\pi+5,t}$ is 1.

There is no need to represent explicitly the shocks that are unobservable. By assuming that the production is perfectly observed after the harvest, it implies that the unobservable part of the seasonal shocks is shifted to the harvest season when it becomes observable.

Let denote $\hat{Q}^r_{i,t}$ the region-r expected production in period $\{i,t\}$, for $i^r_\pi + 1 \le i \le i^r_\pi + 5$, as the product of planned production and realized production shocks:

$$\hat{Q}^r_{i,t} = \mathrm{E}_{i,t}\left[Q^r_{i^r_\pi,t}\exp\left(\sum_{j=1}^{5}\eta^r_{i^r_\pi+j,t}\right)\right] = Q^r_{i^r_\pi,t}\exp\left(\sum_{j=1}^{i}\eta^r_{i^r_\pi+j,t}\right). \tag{5}$$

2.5 Market equilibrium

Markets clear by equalizing the availability, noted $A_{i,t}$, the sum of recently produced and stored commodities to final and storage demand:

$$A_{i,t} = D\left(P_{i,t}\right) + S_{i,t}. \tag{6}$$

2.6 Recursive equilibrium

The variables defining the state of the system change with the season. The availability is always part of the state variables and is defined by available stocks plus current production, if the season is a harvest season:

$$A_{i,t} = \begin{cases} S_{i-1,t} + \epsilon^r_{i,t} Q^r_{i-5,t} & \text{if } i = i^r_\pi + 5, \\ S_{i-1,t} & \text{if } i \neq i^r_\pi + 5. \end{cases} \tag{7}$$

In seasons after planting but before harvesting, defined by the set $\mathcal{I}_\pi = \{i^r_\pi + j | j = \{1, \dots, 4\}$ and $r = \{\text{US}, \text{LAC}\}\}$, the expected production, $\hat{Q}^r_{i,t}$, is also a state variable since it will affect resource allocation via storage. Past storage and the expected production plays a similar role of characterizing the market supply, but they can only be summed at harvest time. Until then, expected production is still uncertain and so not equivalent to stocks. In addition, even if expected production was certain, a future production cannot alleviate a current market scarcity, so the non-negativity of stocks would create a difference between these two state variables.

Let define $s_{i,t}$ the set of state variables at season i and year t. It includes availability and, if relevant, expected production: $s_{i,t} = \left\{A_{i,t}, \hat{Q}^r_{i,t}\right\}$ if $i \in \mathcal{I}_\pi$ and $s_{i,t} = \left\{A_{i,t}\right\}$ if $i \notin \mathcal{I}_\pi$.

From the above we can define the recursive equilibrium of the problem:

Definition. A recursive equilibrium is a set of functions $\mathcal{S}_i\left(s_{i,t}\right)$, $Q^r_{i^r_\pi}\left(s_{i^r_\pi,t}\right)$, and $\mathcal{P}_i\left(s_{i,t}\right)$, defining storage, production, and price over the state variables and the transition equations (5) and (7) such that (i) storer solves (1), (ii) producer solves (2), and (iii) the market clears.

2.7 Welfare

To assess the value of public information, we define welfare as follows. A standard measure of instantaneous welfare, $w_{i,t}$, is provided by the sum of consumers' surplus, producers' surplus, and storers' surplus. But since planting and harvesting occur at different season, the producers' surplus has to be split between revenues and production costs which have to be allocated to their respective season:

$$w_{i,t} = \int_{P_{i,t}}^{P_{\max}} D\left(p\right) \mathrm{d}\, p + P_{i,t} S_{i,t-1} - \left(k\bar{P} + P_{i,t}\right) S_{i,t} + \begin{cases} P_{i,t}\epsilon^r_{i,t} Q^r_{i-5,t} & \text{if } i = i^r_\pi + 5, \\ -\psi\left(Q^r_{i,t}\right) & \text{if } i = i^r_\pi, \\ 0 & \text{if } i \notin \left\{i^r_\pi, i^r_\pi + 5\right\}, \end{cases} \tag{8}$$

where $P_{\max}$ is the maximum price which, since independent from all choices and variables, does not affect welfare comparison. Using equation (7), this expression can be simplified to

$$w_{i,t} = \underbrace{\int_{P_{i,t}}^{P_{\max}} D(p)\,\mathrm{d}\,p + P_{i,t}\left(A_{i,t} - S_{i,t}\right)}_{\text{Consumers' efficiency gains}} - k\bar{P}S_{i,t} - 1_{i=i_\pi^r}\psi\left(Q_{i,t}^r\right). \tag{9}$$

This equation presents three easily interpretable efficiency terms. Combining consumers' surplus with the value of demand removes from the consumers' welfare any transfer to storers or producers related to changes in mean price, leaving only efficiency terms. Then, there are the storage costs, and finally the production costs. This decomposition will prove helpful to understand the source of welfare changes.

Using this expression of instantaneous welfare, we can calculate the intertemporal welfare, normalized to a monthly value by

$$W_{i,t} = (1-\beta)\,w_{i,t} + \beta\,\mathrm{E}_{i,t}\,W_{i+1,t}. \tag{10}$$

The welfare is a function of the season and of the state variables. To remove the dependence on the state variables, we calculate the welfare as an average over the asymptotic distribution of the state variables. Once averaged, welfare varies little with the seasons because of the low discounting we adopt so, without any consequences on the results, we choose the first season to calculate the welfare results.

3 Taking the model to the data

We need to assume functional forms to take the model to the data. We assume functional forms to be isoelastic, with demand given by

$$D(P) = \frac{\bar{D}}{12}\left(\frac{P}{\bar{P}}\right)^{\alpha^D}, \tag{11}$$

where $\bar{D}$ is the steady-state annual demand and $\alpha^D < 0$ is the price elasticity of demand. Similarly, production costs are given by

$$\Psi^r\left(Q^r\right) = \frac{\beta^5\bar{P}}{\left(\theta^r\bar{D}\right)^{1/\alpha^Q}}\frac{\left(Q^r\right)^{1+1/\alpha^Q}}{1+1/\alpha^Q}, \tag{12}$$

where $0 \leq \theta^r \leq 1$ is the steady-state share of region r in world production and $\alpha^Q > 0$ is the supply elasticity, assumed the same in all regions. With this expression of production costs, equation (3) simplifies to

$$Q_{i_\pi^r,t}^r = \theta^r\bar{D}\left[\mathrm{E}_{i_\pi^r,t}\left(\frac{P_{i_\pi^r+5,t}}{\bar{P}}\epsilon_{i_\pi^r+5,t}^r\right)\right]^{\alpha^Q}. \tag{13}$$

$\bar{P}$ and $\bar{D}$ are defined as steady-state values of price and demand in the annual model corresponding to the seasonal model with all decisions collapsed in one period and inter-annual storage and production decision (see section A in the appendix). It is easier to work with these two annual steady-state values rather than with the 24 seasonal values.

Using these isoelastic functional forms, in all the equations the variables in levels can be substituted by the corresponding ratio to annual steady-state levels except for $S_{i,t}/\bar{D}$ and $A_{i,t}/\bar{D}$, which can be interpreted as the stock and availability to annual steady-state use ratios. Similarly, welfare can be expressed unit-free by normalizing $W_{i,t}$ by the monthly steady-state consumption value, $\bar{P}\bar{D}/12$. Since $\bar{P}$ and $\bar{D}$ serve only to define the level of prices and quantities without affecting any of the other results, we normalize their values to 1 and 12. So the monthly steady-state values of price and demand will be close to 1. Table 1 gives the calibration values of all the parameters.

We fix the discount factor by assuming an annual real interest rate of 2%, which is close to the world average preceding the 2008 crisis and the subsequent very accommodating monetary policies (IMF, 2014, Ch. 3). The share of production from the United States, θ^{US}, is 42%, its average 2007–2016 value (considering as only producers Argentina, Brazil, and the United States).

Table 1: Parameterization

Parameter	Economic interpretation	Value	
$\bar{P}$	Steady-state annual price	1	
$\bar{D}$	Steady-state annual demand	12	
β	Monthly discount factor	0.9984	
k	Monthly unit storage cost (% of $\bar{P}$)	0.67	
θ^{US}	Share of US production in total production (%)	42	
α^D	Demand elasticity	−0.2	
α^Q	Supply elasticity	0.4	
		US	LAC
$\sigma^r_{i^r_\pi+1}$	Standard deviation of the 1st news shocks (%)	3.65	2.96
$\sigma^r_{i^r_\pi+2}$	Standard deviation of the 2nd news shocks (%)	4.31	3.21
$\sigma^r_{i^r_\pi+3}$	Standard deviation of the 3rd news shocks (%)	3.63	2.29
$\sigma^r_{i^r_\pi+4}$	Standard deviation of the 4th news shocks (%)	3.33	2.56
$\sigma^r_{i^r_\pi+5}$	Standard deviation of the 5th news shocks (%)	1.73	2.43
$\sqrt{\sum_{i=1}^5 \left(\sigma^r_{i^r_\pi+i,t}\right)^2}$	Standard deviation of the total production shock (%)	7.69	6.07

3.1 Behavioral parameters

This storage model is too complex and too slow to solve to be amenable to a structural estimation. Instead, behavioral parameters are estimated separately, when possible, by using the restrictions implied by their corresponding equations.

3.1.1 Storage cost

To estimate the storage cost, k, we rely on the storage supply curve of Working (1949), which relates the spread between one futures and the spot price to the stored amount. The storage supply curve implies a richer theory of storage than the one used in this paper. Here, we consider only speculative storage: if the spread between nearby and distant futures is not enough to cover storage and opportunity costs, stocks are not carried out. There are actually other motives for stockpiling (e.g., motives of transaction or precaution), so discretionary stocks are never zero and stocks may be carried out at apparent losses. The benefits, other than speculative, that storers derive from holding the physical commodity are dubbed the "convenience yield" (Kaldor, 1939). To distinguish the convenience yield from the storage cost, we assume the convenience yield is almost zero when stocks are largely available, as quantities can be procured easily. Under this assumption, storage costs can be identified from the storage supply curve at high stock levels.

The official marketing year for soybean in the U.S. is from September to August, so available annual U.S. stock data correspond to September levels. For price incentives to be compatible with the stock data, we consider the spread between a distant futures contract and the September spot price. We represent the spread as a percentage of the spot price to remove the need for detrending prices that are unlikely to be stationary across the sample period. Similarly, we normalize stock levels by using the stock-to-use ratio (SUR). We approximate the convenience yield as a function of the inverse of the stock-to-use ratio.[4] This choice ensures that the convenience yield is bounded from above at zero. The storage cost can then be retrieved from the constant of the following linear equation:

$$\frac{F_{i,t|i+n,t}/\left(1+r_{i,t|i+n,t}\right)-P_{i,t}}{P_{i,t}} = nk + b\left(1/\mathrm{SUR}\right), \tag{14}$$

where n is the number of months between the distant futures under consideration and September, $F_{i,t|i+n,t}$ is the distant futures price observed in September, and $r_{i,t|i+n,t}$ is the n-month real interest rate in September. This empirical strategy bears some resemblance to Paul's (1970) one. Paul (1970) estimates storage cost by calculating the maximum spread across different grains, assuming various grains have the same storage cost. Here, we estimate the maximum spread for soybean only but over time.

Prices are from the Chicago Board of Trade. The spot price is the average settle price of the September futures from September 1 to September 14 (the contract expires the 15$^{\text{th}}$). For robustness, we consider two distant futures contracts: January and March with 4 and 6 months horizon, respectively. The distant futures prices are calculated as the average settle prices September 1 to September 14. A n-month nominal interest rate is calculated by interpolating linearly the September secondary market rate of the 3-month and 6-month treasury bills. The corresponding rate of inflation is calculated using the growth rate between September and the distant month of the consumer price index from the Bureau of Labor Statistics. The stock-to-use ratio is calculated by dividing the "Beginning Stocks" by the "Domestic Consumption" for the United States from

[4] A functional form also adopted in Brennan (1958).

PSD (USDA, 2017a). Our sample extends from 1964, when data on stock levels become available, to 2016.

Figure 1 draws the spread with respect to the stock-to-use ratio. As expected, spreads tend to be higher for the March horizon than for January, as storage cost increases with the longer horizon. We can observe the standard shape of a storage supply curves with mostly positive spreads for high stock levels, and negative ones for low stock levels. Figure 1 draws also the storage supply curves estimated from equation (14) using ordinary least squares, with estimates available in table 2. The shape of the storage supply curves is the same for both horizon, since the estimated coefficients on the inverse of stock-to-use ratio are almost identical. The storage curves differ only by their constant, which would tend to confirm its interpretation as the physical storage cost. When the constant is corrected for the number of months, one obtains the same storage cost from the two estimates: $k = 0.0067$. We retain this value for the simulations. On an annual basis, this result implies an annual storage cost of 8.04% of the steady-state price.

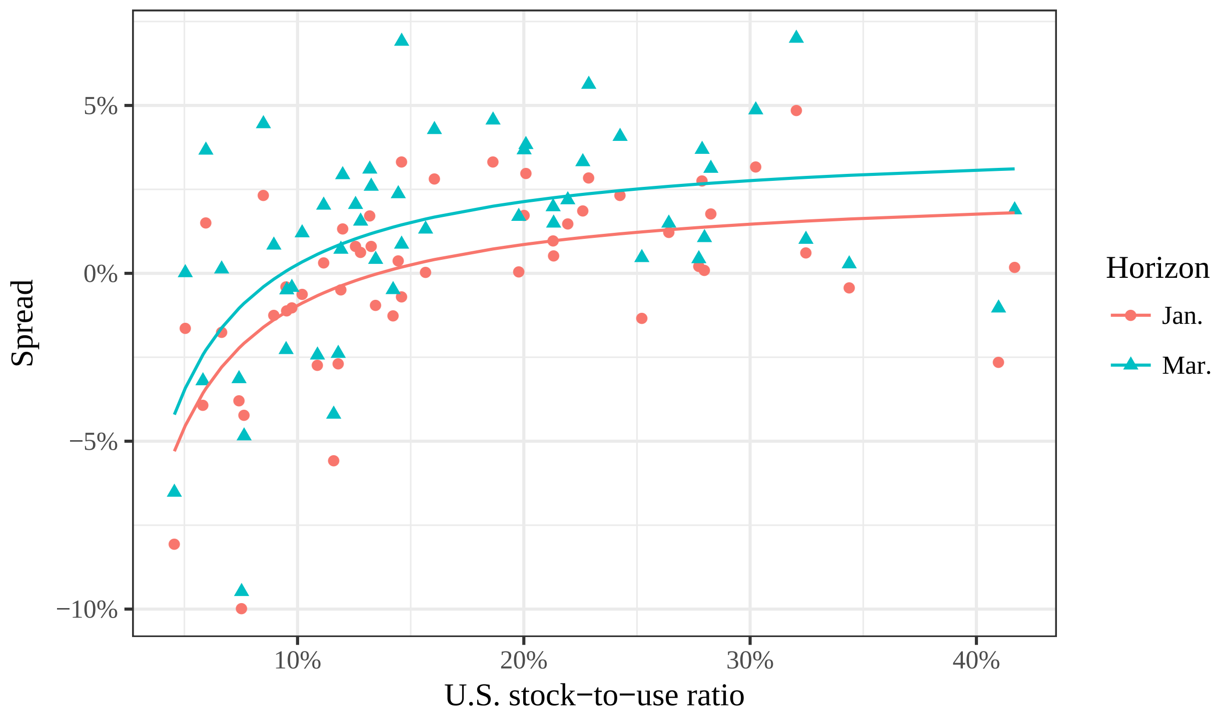

Figure 1: Interest-adjusted price spread vs. inventory and estimated storage supply curves

Based on the same decomposition between storage cost and convenience yield, another estimation approach could have been possible. By taking the difference between the spreads at different horizon, we can purge the data from the convenience yield and we are left with the storage cost for the remaining 2-month period between January and March. The resulting difference of spreads is not correlated with the stock-to-use ratio (figure A1 in appendix). Divided by 2, the number of months, its mean value is 0.0062 and its median 0.0068, which confirms the previous estimation of storage cost.

Table 2: Estimates of storage supply curve

	January	March
Constant (nk)	0.0268***	0.0401***
	(0.0062)	(0.0074)
Inverse of U.S. stock-to-use ratio (b)	−0.0036***	−0.0037***
	(0.0007)	(0.0008)
R^2	0.3502	0.2870
Observations	53	53

Note: *** indicates significance at the 99%.

3.1.2 Supply elasticity

To estimate the supply elasticity, we adopt Roberts and Schlenker's (2013) instrumental variable approach that we modify to fit our framework. With respect to the estimation of supply elasticities, Roberts and Schlenker's key insight is twofold. First, they argue that using futures price as explanatory variable is not enough to purge the model from endogeneity because they can reflect anticipated shocks that are unobserved to the econometrician. Second, because speculative storage links prices between periods, futures prices are affected by past shocks, and so the past shocks can be used as instruments.

We do the following adjustments to Roberts and Schlenker's empirical model. Since we are interested in the supply response of the United States and Argentina and Brazil combined, and the two regions do not face the same price incentives because of their crop calendars, we do not estimate like these authors the model on a world aggregate. We create a panel with the two regions. Our dependent variable is the log quantity supplied, denoted by $q_{rt} = \log\left(Q^r_{i^r_\pi+5,t}\epsilon^r_{i^r_\pi+5,t}\right)$. Our explanatory variables are the log of harvest-time price as expected at planting time, $p^q_{rt} = \log\left(\mathrm{E}_{i^r_\pi,t} P_{i^r_\pi+5,t}\right)$, corresponding empirically to futures price, and the region-specific random shock, $\omega_{rt} = \epsilon^r_{i^r_\pi+5,t}$. The random shock is included as an explanatory variable to purge the dependent variable from its stochastic component and to reduce the endogeneity problem as the realized shocks are likely to be a good proxy for anticipated supply shocks (Hendricks et al., 2015).

For the first-stage equation, we follow Roberts and Schlenker (2013) and use as instrument lagged yield shocks, but to be consistent with our theoretical model we distinguish the yield shocks by Hemisphere.[5] We denote ω_{ht} the yield shock in the Hemisphere h. The most recent shock to instrument the futures price faced by U.S. farmers is the Southern Hemisphere harvest, which is actually contemporaneous to planting in the U.S. Similarly, the Northern Hemisphere shock is used to instrument the futures price faced by farmers in Argentina and Brazil. In our case, lagged supply shocks alone tend to be a weak instrument. Indeed, theoretically, because of storage, supply shocks have a time-varying role in determining prices: they have a limited effect on prices when stocks are abundant and, conversely, an important one when stocks are low. So, we use also an additional instrument, not used in Roberts and Schlenker (2013). To account for the

[5] See also Winne and Peersman (2016) for a recent work using the timing of harvests around the world to build exogenous variables.

fact that lagged yield shocks have a differential effect depending on the stock levels, we add as instrument lagged harvest-time prices, $p^d_{rt-1} = \log\left(P_{i^r_\pi+5,t-1}\right)$, and their interaction with the hemispheric yield shocks. Harvest-time prices can be seen as a proxy for market availability and its interaction with yield shocks will account for the possibly non-stationary effect of yield shocks.

The resulting empirical model is

$$q_{rt} = \iota + \alpha^Q p^q_{rt} + \gamma\omega_{rt} + f^q_r(t) + \delta_r + u_{rt}, \tag{15}$$

$$p^q_{rt} = \zeta + \kappa\omega_{rt} + \lambda\omega_{ht-1} + \nu p^d_{rt-1} + \xi\omega_{ht-1}p^d_{rt-1} + f^p_r(t) + \tau_r + v_{rt} \text{ for } r \notin h. \tag{16}$$

Each equation includes region-specific time trend $f^j_r(t)$, modeled by restricted cubic spline ($j \in \{q, p\}$),[6] and region fixed effects, δ_r and τ_r. This empirical model is fully consistent with our theoretical model except for the anticipated shocks at planting time (unobserved by the econometrician) that justify using an instrumental variable approach and that we neglect for simplicity in our theoretical model.

Soybean production and yields are from FAOSTAT (FAO, 2017). Countries are classified into Northern and Southern Hemispheres using the crop calendar of Sacks et al. (2010): countries with mean planting date between February 19th and July 19th are classified in Northern Hemisphere. For the countries not covered by Sacks et al. (2010), which concerns mostly minor producing countries, we use a simple heuristic and assign to the Northern Hemisphere all countries with a capital city located above the latitude of Mexico City. Futures prices are from the Chicago Board of Trade with a delivery month of November for the United States and of May for Argentina and Brazil. p^d_{rt} is constructed as the log of the average futures price during the month of delivery and p^q_{rt} as the log of the average futures price in April for the United States and in October of the previous year for Argentina and Brazil. Yield shocks are constructed using the method described in Roberts and Schlenker, so ω_{ht} is the average across countries of Hemisphere h of log yield deviations from a three-knot restricted cubic spline trend. The hemispheric shocks are scaled to represent deviations from the trend world production, so that their sum is identical to the yield shock variable (ω_t) used in Roberts and Schlenker. For example, $\omega_{ht} = -0.02$ indicates a -2% shock to world production related to the deviation of the average yield in hemisphere h from its trend. Similarly, ω_{rt} is the region r log yield deviations from trend, calculated using jackknifed residuals. FAOSTAT data are available from 1961 to 2014, but because of our lagged yield shocks our sample starts in 1962.

Estimation results are presented in table 3 and include 2SLS and OLS estimates with different specifications of the time trend. The first stage is strong, as indicated by the F-statistics. The overidentification test does not reject the validity of the instruments. The first-stage coefficients follow the intuitions from theory. The coefficient on p^d_{rt-1}, which can be interpreted directly because ω_{ht} has zero mean, indicates that lagged delivery prices influence positively harvest-time prices consistently with storage theory. As expected, the

[6] Following the heuristics proposed in Harrell (2001), the knots for the cubic spline with three knots are located in 1967, 1988, and 2009. The spline with four knots uses as knots 1965, 1980, 1996, and 2011 and the spline with five knots 1965, 1976, 1988, 2000, and 2011.

marginal effect of lagged yield shock, ω_{ht-1}, depends on market condition: evaluated at mean delivery price it is -1.68, at the minimum sample price it is -0.01, and -3.92 at the maximum. The estimated supply elasticities are higher than Roberts and Schlenker's estimate for a caloric aggregate, but similar to their estimate for soybean alone in a four-crop system (Table A7). They are also similar to Haile et al.'s (2016) estimates. The estimates are sensitive to the flexibility of the time controls, with a much higher supply elasticity for a three-knot spline (in 2SLS and OLS).[7] This sensitivity, which was not present in Roberts and Schlenker (2013) is explained by the fact that our dependent variable is the production of regions and not of the world. World production has a linear trend over the period and so is easy to detrend, while the soybean production of Latin American countries took off only in the 1970s after the 1973 U.S. embargo on the export of soybeans and so requires more flexibility in the trend. The supply elasticities estimated by OLS are much lower than under 2SLS and the Hausman test confirms the endogeneity of futures price. These differences between 2SLS and OLS illustrate also that Hendricks et al.'s (2015) conclusion that it is not necessary to instrument futures price when current yield shock is included in the explanatory variables does not generally hold. Consequently, our preferred estimate uses the 2SLS estimator and the more flexible time trend and we retain for the simulations a supply elasticity of 0.4.

3.1.3 Demand elasticity

We have not been able to estimate the demand elasticity using Roberts and Schlenker's method. In their paper, they obtain a statistically significant demand elasticity when estimating a four-crop system only by imposing symmetry and by using USDA data rather than FAOSTAT data (Table 8). Their elasticity is -0.236. We do not know any other recent source of credibly estimated demand elasticity for soybean, so we follow Roberts and Schlenker and retain a demand elasticity of -0.2.

3.2 Seasonal shocks

To calibrate the size of the seasonal shocks, we use the monthly changes in production projections from the World Agricultural Supply and Demand Estimates (WASDE) reports (USDA, 2017b). To make the link between the model and the data, we can note that using equation (5), we have $\exp\left(\eta^r_{i,t}\right) = \hat{Q}_{i,t}/\hat{Q}_{i-1,t}$: the seasonal shocks can be identified by calculating the ratio between two consecutive projections. Given the assumption of shocks with unitary mean, we have $\operatorname{var}\left(\hat{Q}_{i,t}/\hat{Q}_{i-1,t}\right) = \exp\left(\left(\sigma^r_i\right)^2\right) - 1 \approx \left(\sigma^r_i\right)^2$. One benefit of the assumption of multiplicative shocks is that there is no need to detrend the data to compare shocks across years despite the trending production level. From these seasonal shocks, table 4 reports the mean and standard deviation of the monthly adjustments to production projections. For all months, we cannot reject the null hypothesis that the mean is equal to 1, as assumed in the model. So, the standard deviation can be interpreted as a coefficient of variation and have been multiplied by 100 for ease in interpretation.

[7] With a six-knot spline the estimation results are very close to the five-knot spline and so have not been reported.

Table 3: Estimates of supply elasticity

	2SLS			OLS		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
Supply elasticity α^Q	0.60^{***}	0.25^{**}	0.38^{***}	0.51^{***}	0.09	0.15^{**}
	(0.08)	(0.10)	(0.10)	(0.07)	(0.07)	(0.07)
Shock ω_{rt}	1.60^{***}	1.39^{***}	1.21^{***}	1.62^{***}	1.41^{***}	1.24^{***}
	(0.22)	(0.15)	(0.15)	(0.21)	(0.15)	(0.14)
First stage ω_{rt}	0.16	0.10	0.08			
	(0.16)	(0.14)	(0.15)			
First stage ω_{ht-1}	10.48^{*}	9.68^{*}	10.08^{*}			
	(6.00)	(5.45)	(5.50)			
First stage p^d_{rt-1}	0.79^{***}	0.58^{***}	0.57^{***}			
	(0.04)	(0.06)	(0.06)			
First stage $\omega_{ht-1}p^d_{rt-1}$	-2.00^{**}	-1.84^{**}	-1.91^{**}			
	(0.97)	(0.88)	(0.89)			
First-stage F-statistics	116.03	38.77	35.67			
p-value for Hausman test	0.02	0.01	0.00			
p-value for overid. test	0.55	0.77	0.79			
Observations	106	106	106	106	106	106
Spline knots	3	4	5	3	4	5

Note: The first three columns, 1a–1c, use two-stage least squares, while columns 2a–2c use ordinary least squares. Columns a, b, and c include restricted cubic splines in time with three, four, and five knots, respectively. Coefficients in the first two rows give the results for log supply, while coefficients from the third to the sixth row give first-stage results of log price. Coefficients on time trends and regions are suppressed. ***, **, and * indicate significance at the 99%, 95%, and 90% levels, respectively.

For the U.S. market, the projections for the new harvest begin with planting in the reports of May (in earlier reports, the projections for the new crop year started later in June or July), but without any changes in projections between May and June. The usual planting and harvesting dates for U.S. soybean are as follows (USDA, 2010). Planting of soybean crops begins in late April, early May and ends in early July, with the most active period being in the second half of May. The U.S. soybean harvest begins in September and most of it is finished by the end of November. This is consistent with what we observe in WASDE reports. Most of the production changes occur between June and November, with the largest changes being observed in the August report which is the first to include survey-based yield forecast, not just trend yields. Between November and December, the production projections are not adjusted. Small adjustments occur between December and March because of data revisions. Further revisions may occur several months later, when additional data (e.g., crushing or exports) become available.

Based on the usual planting and harvesting dates and the WASDE reports, we assume that the first season in the model, $i = 1$, when U.S. soybeans are harvested correspond to the period running from October 15 to November 15, noted October/November. So, we represent the U.S. soybean crop cycle in the model as

Table 4: Mean and standard deviation of monthly adjustments to production projections

	U.S. (1973–2016)[a]			Argentina and Brazil (1984–2016)[a]					
					Argentina		Brazil		LAC[b]
Month	# Obs.	Mean	SD (%)	# Obs.[d]	Mean	SD (%)	Mean	SD (%)	SD (%)
June	22	1	0	8	1	0	1	0	0
July	30	0.992	3.654	13	0.999	0.509	1	0	0.199
Aug.	42	0.991	4.309	23	1.002	0.753	1.003	1.189	0.708
Sept.	43	0.996	3.625	23	1.007	1.214	1.001	1.523	1.015
Oct.	43	1.004	3.325	31	1.007	2.011	1.006	2.086	1.607
Nov.	43	1.006	1.730	31	1.006	1.705	1.004	1.610	1.347
Dec.	39	1	0	32	1.005	1.426	0.999	0.839	0.660
Jan.	39	1.001	1.382	33	1.001	1.624	1.000	3.234	2.313
Feb.	39	0.999	0.405	33	1.000	4.836	1.002	2.322	2.508
Mar.	39	1.000	0.009	33	1.003	2.194	1.004	2.493	1.784
Apr.	44	1	0	33	1.001	2.794	0.997	2.285	1.998
May	39	1	0	32	0.988	2.987	0.999	2.246	1.899
June	39	1	0	32	1.001	2.903	1.002	1.365	1.139
July	43	1.000	0.267	32	0.995	2.258	1.000	0.975	0.906
Aug.	43	1	0	32	0.999	1.145	1.002	0.541	0.544
Sept.	43	1	0	23	1.001	0.887	1.000	0.378	0.396
Oct.	42	1.002	1.456	22	1.001	0.461	1.003	0.722	0.555
Nov.	42	1.000	0.009	22	1.002	1.242	1.001	0.407	0.515
Between-year volatility[c]			14.791			21.656		12.844	11.674

Source: Authors' calculation based on WASDE (USDA, 2017b) and PSD (USDA, 2017a). When there was more than one production forecast per month, as in the 1970s, we selected the one with the release date closest to the tenth of the month.
Note: [a] Marketing years used to calculate the statistics. [b] The last column contains the standard deviation of monthly adjustments to production projections of Argentina plus Brazil. [c] Calculated as the standard deviation of the first-order difference of the logarithm of the final production as reported in PSD for the same marketing years as the monthly adjustments. [d] There are fewer observations outside of the months of interest for the Latin American countries, because some could not be collected automatically from the original pdf files. Given they would not affect any of the results, they have not been collected manually.

starting with the planting decision in May/June and ending with harvesting in October/November. There are five months after planting and a seasonal shock in each. The standard deviations of the seasonal shocks, σ_i^{US}, are simply the standard deviations of the production adjustments between July and November in table 4. Using these five standard deviations for the seasonal shocks implies a standard deviation for the equivalent aggregate shock of 7.7%. This number can be compared to a between-year volatility of production of 14.8% (standard deviation of the first-order difference of the logarithm of production values in PSD): the combined seasonal shocks appear to account for a sizable share of this volatility. The rest of production volatility can be attributed to shocks occurring before planting, to the endogenous reaction of farmers to market incentives, and to unobservable seasonal shocks that are only revealed after the harvest in the data revisions.

Our modeling framework neglects two types of shocks related to the production. First, it neglects the possibility of a shock occurring before planting and observable by farmers and other agents, which could be related for example to the pre-plant precipitation or groundwater levels (an issue recently addressed in

Huang and Moore, forthcoming). There is no source of information that could allow us to recover directly the size of such a shock. Second, it neglects the revisions made to the production estimates a few months after the harvest. Since soybean is mostly crushed or exported, its use can be accounted with more precision than for other crops with more dispersed uses (such as maize or wheat). The data on use provide information on the size of the recent harvest that are used to update the production estimates. In addition, it takes a couple of months for the collection of information about final yields to be complete. These two pieces of information explain the non-zero numbers for January, February, and for later months after the harvest in table 4. These data revisions involve a different type of informational problem than the one studied in the present article. They imply that the state of the system is observed only with measurement errors and the new observables at each period allow the agents to update their estimation of the state. Such a problem relates more to issues of learning—how agents extract information about the state from observables, and how it affects market outcomes (studied recently in the context of the oil market in Leduc et al., 2016, and in the context of revisions to macroeconomic data in Galvão, 2017)—than to a problem of news shocks. So, given our purpose of assessing the economic value of news shocks, neglecting these post-harvest shocks should not affect our results.

For Argentina and Brazil, the identification of the seasonal shock structure is more challenging. Taken together, these countries present a much larger agro-climatical diversity than the United States, so the planting and harvesting of the main soybean crops is more spread in time. In addition, because of their tropical climate, double- or triple-cropping systems are possible, spreading soybean production even further in time. The WASDE reports started to cover Argentina and Brazil, separately, only in January 1985, with 1982 being the first marketing year with some data and 1985 the first marketing year with complete data, therefore we have fewer observations than for the U.S. The projections of production for these countries present different patterns than those for the U.S. They start with the U.S. projections, in May for the most recent reports, so several months before planting in the Southern Hemisphere. The initial projections are estimates based on trends by USDA oilseeds experts of the likely production levels. These early projections may be adjusted before the planting seasons, based on additional information. The variations between May and October are different from pure news shocks as they are the difference between the USDA trend estimates and the first information on planting conditions or planting intentions. Another different feature, compared to the U.S. projections, is that these countries projections do not seem to converge smoothly to some final estimates, even in August, months after the end of the main harvest. Crop adjustments are much larger in Argentina than in Brazil or the United States, but because it is a smaller producer than Brazil and their production shocks are not correlated, the volatility of the region composed of these two countries is often smaller than the volatility of the countries taken separately.

In order to obtain for this region a series of five seasonal shocks, we must make some assumptions. Based on AMIS crop calendar,[8] October and November are the two months when most soybean planting is done

[8] http://www.amis-outlook.org/amis-about/calendars/en/, retrieved March 27, 2017.

in Brazil, and the harvest takes place in March and April. In Argentina, planting occurs in November and December and harvesting in April and May. In the model, planting for the region is assumed to take place in October/November and harvesting in March/April. We consider that the main news shocks are the five shocks with the largest standard deviation between January and May. But if we were to use directly the values for these months in the last column of table 4, this would result in a standard deviation of the total production shock of 4.74%, much below what is observed inter-annually at 11.67%. Consequently, we rescale the seasonal shocks so that the total production shock represents 52% of the inter-annual volatility, what is observed for the United States. The resulting standard deviations are available in table 1. They display a variability that is more regularly spread than for the United States.

4 Quantitative analysis

We now turn to the analysis of the effects of news shocks. To do this, we compare the model with news shocks described above with a model without news shocks where the production shocks are degenerate and concentrated in the harvesting periods. Table 5 presents descriptive statistics on the asymptotic distribution of prices and stocks for the models with and without news shocks.[9] The changes in mean prices are not represented in the table because these changes are very small (below 0.5%). News shocks affect the distribution of prices without changing their mean.

Before analyzing the effects of introducing news shocks, let us first explain the pattern of seasonal price volatility in the absence of these shocks. Without news shocks, the standard deviation of seasonal prices varies with the seasons, but not much (2.08 percentage point maximum), increasing by about 0.03 percentage point between seasons except for new harvests, when there are small jumps. These variations can be easily explained as follows. Between one harvest and the last month before the next one, there is no new information and stocks are always positive, so from equation (1) successive prices are deterministic functions of the harvest-time price: $P_{i+1,t} = \left(P_{i,t} + k\bar{P}\right)/\beta$. If the only storage cost was the opportunity cost, the standard deviation and the mean price would increase between two periods by $1/\beta$, but the coefficient of variation would stay constant. The presence of additive storage costs breaks this multiplicative relationship and the coefficient of variation slightly changes with period.

The presence of news shocks changes a lot the seasonal price volatility. The differences between seasons become much larger with a maximum spread of 3.81 percentage point. In most seasons the volatility decreases significantly, but not in February/March, July/August, August/September, and September/October, with the volatility increase being the highest the closest to the harvests (plus 14 percentage point in September/October before the U.S. harvest). As expected, between the seasons when no news arrives, there is the same deterministic increase of 0.03 percentage point in price volatility as observed in the model without any news shocks. The decrease in price volatility observed in the other seasons follows from a simple intuition: having

[9] See section B in the appendix for details about the numerical methods.

Table 5: Descriptive statistics on the asymptotic distribution (all results in percentage)

		Prices					Stocks
		Standard deviation			Skewness	Kurtosis	Mean
Period	Shock	No news	News	Changes	Changes	Changes	Changes
Oct./Nov.	σ_1^{US}	18.27	15.47	−15.4	−26.1	−41.9	−1.3
Nov./Dec.	σ_2^{LAC}	18.30	16.10	−12.0	−34.9	−49.8	−1.7
Dec./Jan.	σ_3^{LAC}	18.33	17.04	−7.1	−43.4	−58.4	−2.5
Jan./Feb.	σ_4^{LAC}	18.36	17.65	−3.9	−46.7	−62.2	−4.2
Feb./Mar.	σ_5^{LAC}	18.45	18.63	1.0	−48.6	−65.5	−15.7
Mar./Apr.	σ_6^{LAC}	16.37	15.17	−7.4	−19.6	−30.8	−0.9
Apr./May		16.40	15.19	−7.4	−19.5	−30.8	−1.0
May/June		16.43	15.22	−7.4	−19.6	−30.8	−1.3
June/July	σ_9^{US}	16.45	15.75	−4.3	−27.5	−38.4	−1.6
July/Aug.	σ_{10}^{US}	16.48	16.74	1.6	−35.6	−48.1	−2.3
Aug./Sept.	σ_{11}^{US}	16.51	17.74	7.5	−37.4	−51.5	−4.1
Sept./Oct.	σ_{12}^{US}	16.64	18.98	14.1	−33.3	−48.6	−16.7
Inter-annual[a]		13.41	13.52	0.8	−25.1	−32.1	–
Intra-annual[b]		3.69	5.00	35.3	–	–	–

Source: Statistics calculated over 1,000,000 sample observations from the asymptotic distribution simulated with the model.
Note: [a] Annual price calculated as $\sum_{1=1}^{12} P_{i,t}/12$.
[b] Intra-annual standard deviation calculated as $\mathrm{E}\sqrt{\sum_{i=2}^{12}\left[\ln\left(P_{i,t}/P_{i-1,t}\right)-\ln\left(P_{12,t}/P_{1,t}\right)/11\right]^2/10}$.

advance information allows the market to adjust stock levels before the realization of the harvests, which avoid abrupt changes at harvest time; adjustments are smoothed over the growing seasons. But the increased volatility just before the harvest is less intuitive. Its explanation is that most of the aggregate productive shock has been observed before the harvest (over the four months of the growing season before harvest, the news shocks aggregate to standard deviations of 7.49% for the US and 5.56% for LAC). So even if the harvest is one period after, the market is already fairly confident about its size. On the other hand, the seasons just before harvests are also the season when stocks are at their lowest levels, so seasons when the market is very sensitive to news because adjustment capacities are limited. That the information about the harvest size has been moved to seasons when the market is less able to cope with shocks explains the increased volatility in these seasons.

The most dramatic effects of the introduction of news shocks can be observed in the higher-order moments: skewness and kurtosis. In all seasons, as expected with the presence of speculative storage, prices are positively skewed (not represented in table 5): prices are concentrated around below the mean with occasional positive deviations that are possibly larger than negative ones. News shocks reduce a lot the skewness, between 20 and 49%, by reducing significantly the occurrence of high prices. Similarly, the kurtosis of the price distribution is at least reduced by one-third (although the distribution remains leptokurtic) by the introduction of news shocks, because of the decreased probability of extreme price events.

A striking result is that this reduction in seasonal price volatility and in the occurrence of high prices is obtained despite a reduction in mean stock levels. Without news shocks, it is profitable to keep higher level of stocks before the harvest because there is a non-negligible likelihood of a bad harvest, so keeping more stocks might be profitable. With news shocks, before the harvest the market has a pretty good idea of the coming harvest so the speculative motive is decreased. The reduction in stocks level in periods that are not just before harvests seems less impressive because in these periods the mean stock levels are much higher just for smoothing the bi-annual harvest over the whole year. But in levels, and not in proportion, the decrease is almost the same for all periods. Stocks are not always reduced with news shocks. If the early news shocks indicate a bad harvest, destocking is reduced compared to a situation without news which mitigates the price increase.

Table 5 presents also statistics on the inter- and intra-annual price volatility. Inter-annual price volatility is calculated by averaging seasonal prices over the U.S. marketing year. This approach is similar to what is done for the World Bank Pink Sheets, except that the World Bank averages over the calendar year. The inter-annual price volatility is significantly lower than any of the seasonal price volatility, because averaging over the seasons removes the intra-annual volatility, related to the fact that the crop cycles do not overlap and so producers can react to the observation of the other Hemisphere harvest.[10] The size of the inter-hemispheric supply response was recently analyzed in Lybbert et al. (2014), but the present article is the first to quantify its effect on price volatility. The level of inter-annual price volatility is barely affected by the presence of news shocks, despite the important changes to seasonal price volatility. This shows that one of the main effects of news shocks is to reorganize the price volatility throughout the seasons. However, regarding the higher-order moments of the distribution, the inter-annual results are similar to the seasonal ones: very strong reduction in skewness and kurtosis related to a lower occurrence of price spikes. An annual statistic calculated as an average would not make much sense for stocks. Annual stocks are usually considered to be the level at the beginning of the marketing year. Since in our model, the seasons when there are non-zero probabilities of stockouts are Feb./Mar. and Sept./Oct., the change in the annual stocks level could be considered to be the change in one of these seasons statistics.

Intra-annual price volatility is calculated as the mean over the asymptotic distribution of the bias-corrected root mean square of the difference between within-year month-to-month price changes and yearly price changes: $\mathrm{E}\sqrt{\sum_{i=2}^{12}\left[\ln\left(P_{i,t}/P_{i-1,t}\right)-\ln\left(P_{12,t}/P_{1,t}\right)/11\right]^2/10}$. This measure of intra-annual price volatility increases significantly with news shocks. Indeed, in the absence of news shocks, neglecting storage costs, intra-annual volatility is only related to the Latin American harvest, which is the only within-year shock in the model that would justify non-deterministic changes in prices. Outside harvest periods, without news shocks prices are deterministic functions of past prices, and constant without any storage cost. With news shocks,

[10]In the inter-annual model in appendix (section A) where crop cycles are collapsed, price volatility is much higher than the inter-annual price volatility of the model without news because of the lack of adjustments after one harvest.

this is different: except in 2 periods without news shocks there are always some informational shocks and prices adjust more regularly resulting in a higher intra-annual price volatility.

The first-order auto-correlation of the inter-annual price average increases by 7% percentage point when news shocks are introduced (a result also present in Ng and Ruge-Murcia, 2000, and Osborne, 2004). This is an interesting result because it is obtained despite lower stock levels, knowing that higher stock levels are usually associated to higher auto-correlation levels. The intuition behind this result is that news shocks create links between different marketing years given that the conditions of the coming harvest are mostly known a few months before.

To summarize the consequences in terms of market dynamics of news shocks: the presence of news shocks leads to a different repartition of seasonal price volatility, with more volatility just before harvests and less volatility at other seasons, but has little effect on the inter-annual price volatility; it reduces stock levels; it reduces the occurrence of price spikes by allowing stocks to adjust before the harvest.

Table 6 decomposes welfare gains in the three efficiency components identified in equation (9): consumer's efficiency gains, storage costs, and production costs. Having news shocks compared to having only one production shock at the harvest increase welfare. Total gains amount to 0.45‰ of the value of annual steady-state consumption. To better understand the level of the welfare gains, we can compare them to the maximum gains attainable by having advance information about the coming harvests. For this we assume that the total production is fully known one period after planting and that there are no other shocks after. In this case, total welfare gains equal 0.73‰ of the value of annual steady-state consumption. So, the actual news shocks achieve 62% of the potential gains. What explains the small size of the gains, especially compared to the important changes in price dynamics shown in table 5, is the simplicity of the model that does not contain many channels through which price volatility could affect welfare. Consumers are not risk averse, so they are mostly indifferent to price volatility (Gouel, 2014). Similarly, production costs have little reasons to be affected by the news shocks since the production decisions are taken without news anyway and mean prices barely change. The only significant source of welfare gains is related to the storage costs. Because of the advance information, fewer stocks are needed. The reduction in storage costs explains most of the welfare gains but, given that storage costs are small compared to the value of consumption, the eventual welfare gains are also low. That most of the welfare gains are related to the decreased in inter-annual stock levels confirms the choice made in this article of a storage model accounting for both inter- and intra-annual market dynamics, whereas the early approaches of Hayami and Peterson (1972) and Bradford and Kelejian (1977, 1978) focused only on intra-annual storage and so missed most of the effect of news shocks.

Given that most welfare gains are related to reduction in storage costs, it is reasonable to express them in proportion to initial storage costs. In this case, the total welfare gains from news amount to 2.31% of storage costs. But, we should not expect most storage costs to be affected by the presence of news. In a storage model with intra-annual dynamics, most storage costs are incurred to smooth consumption between harvests. These irreducible storage costs can be evaluated from the steady-state values of stock levels, which are determined

Table 6: Welfare changes with respect to a situation without news shocks (‰ of the value of annual steady-state consumption)

Welfare element	Actual news shocks	All news in the first shock
Consumer's efficiency gains	0.08	0.14
Storage costs	0.41	0.65
Production costs	−0.04	−0.06
Total gains	0.45	0.73

by assuming risk away. If we express welfare gains in proportion to "risky" storage costs (storage costs minus their steady-state value), gains represent 18.76% of these costs. So, a sizable portion of the storage costs generated by uncertainty are reduced by the provision of news shocks. These gains are consistent with the reduction in ending stocks observed in table 5.

To illustrate the welfare gains, we use a price of soybean of $350 per ton and a world production of 320 million tons, based on recent years' data. The world production is then valued at $112 billion and the benefits from news shocks at $50.4 million. This value is for the total value of information assuming as counterfactual a situation without advance information at all which overestimates the gains as some information would be collected, at least privately, in the absence of USDA. As a comparison, the combined costs of the three main services involved in the elaboration of the WASDE reports—the National Agricultural Statistics Service, the Foreign Agricultural Service, and the World Agricultural Outlook Board—were $211 million in 2013 (C-FARE, 2016, Table 1) for missions that go beyond this commodity and these monthly reports, and for analyses that include also the demand side of the balance sheets, which we did not consider here.

5 Conclusion

This article has shown, through the example of the global soybean market, how news shocks that bring advance information about future production affect market dynamics compared to a situation where the production is known when it is realized. The results have been obtained by building a rational expectations storage model able to account for intra-annual and inter-annual dynamics and by running counter-factual simulations. The most important parameters of the model have been estimated, when possible, and the monthly USDA reports of crop production forecasts have been used to calibrate realistic news shocks in the model. These shocks matter a lot for market dynamics because based on the information they brought storers will adjust their stocks before the harvest to be more consistent with the newly expected market tightness. These early stocks adjustments reduce the needs to carry large inter-seasonal stocks, these latter being reduced significantly. In addition, when a bad harvest is forecast, storers reduce their destocking before the harvest, which reduces the occurrence of price spikes a lot, even if the overall price volatility is little affected by the presence of news shocks. The reduction in stock levels allowed by news shocks explained most of the welfare gains from having advance information, which amount to 2% of the value of storage costs. These benefits, even if small, are commensurate with the actual public costs of data collection.

This study deals with only a small part of the informational issues in commodity markets, and so with a small part of the potential benefits of public information in these markets. One can cite for example the problem of the reporting of stocks. Stock levels are not directly observable by market participants but are regularly reported by governmental agencies and international organizations. These statistics, while crucial for the market, are also known to be of limited reliability: they are estimated with little precision either based on surveys or, more frequently, as a residual in commodity balances. In the framework of the storage model, it means that agents do not observe the state variables with precision and must make decisions with incomplete information. The reporting of stock levels is likely to be a problem as economically important, if not more, than the advance information about the coming harvests. Another feature of our approach that may have led to underestimate the benefits from public information is that we have developed a simple model with few margins of adjustments to new information. Any extension where agents must take costly decision during the growing season would likely increase the benefits. This could be related for example to international trade costs, planting in crops that are close substitutes, or buying feed for livestock.

Beyond the question of the cost/benefit analysis of public information, this research may have other policy implications in relation to the opportunity of price stabilization policies. Following the 2007/08 food crisis, we have observed a renewed interest in food price stabilization policies (Gouel, 2014) in particular to protect consumers from price spikes. These policies involve often the combination of storage and trade policies that can be very costly fiscally (Gouel and Jean, 2015). The present results show that the public provision of information during the growing season is a good way to decrease the occurrence of price spikes without the need for public interventions in the market. Since market information systems in developing countries are very likely much less developed than in the United States, their development should be considered before price stabilization policies.

The results of this article on seasonal dynamics have also consequences for future research on the annual storage model. They illustrate some challenges for the estimation of this model that have been barely acknowledged in the literature in the tradition of Deaton and Laroque (1992, 1996). Only Guerra et al. (2015) touch on the issue by showing the important differences in estimating an annual model using harvest-time prices or annual averages, which is consistent with our results: the price volatility of the annual average is very different from the volatility in every season. But an even more difficult issue is that calibrated with the same parameters, an annual storage model, which implicitly neglects news shocks, and an intra-annual model with news shocks will present very different price dynamics with much fewer price spikes in the second model because of the possible stock adjustments before harvest. So, the estimation of an annual model on prices generated by a more realistic seasonal model would likely not recover the original structural parameters. It remains to be seen what parameters are most affected by this misspecification.

Appendix

A Inter-annual storage model

We present here the inter-annual storage model corresponding to the seasonal model used in the article. It is defined by four equilibrium equations:

$$S_t : \beta^{12} \mathrm{E}_t\, P_{t+1} - P_t - 12k\bar{P} \leq 0, = 0 \text{ if } S_t > 0, \tag{A1}$$

$$Q_t^r : \beta^{12} \mathrm{E}_t \left(P_{t+1}\epsilon_{t+1}^r\right) = \Psi_I^{r\prime}\left(Q_t^r\right) \text{ for } r \in \{\text{US,LAC}\}, \tag{A2}$$

$$P_t : A_t = 12D\left(P_t\right) + S_t, \tag{A3}$$

and one transition equation:

$$A_t : A_t = S_{t-1} + \sum_{r\in\{\text{US,LAC}\}} \epsilon_t^r Q_{t-1}^r. \tag{A4}$$

A few adjustments have to be made to transform the model to a purely inter-annual model. Since the production decision is taken one year before the harvest, the crop is assumed to grow for one year, involving a discounting different on the left-hand side of equation (A2) compared to the left-hand side of equation (3). Because the discounting is increased, the marginal cost is increased similarly with

$$\Psi_I^r\left(Q^r\right) = \frac{\beta^{12}\bar{P}}{\left(\theta^r\bar{D}\right)^{1/\alpha^Q}} \frac{\left(Q^r\right)^{1+1/\alpha^Q}}{1+1/\alpha^Q}. \tag{A5}$$

It can be easily verified that this model has for steady state: $A^{ss} = \bar{D}$, $S^{ss} = 0$, $Q^{r,ss} = \theta^r\bar{D}$, and $P^{ss} = \bar{P}$.

B Numerical methods[11]

The rational expectations storage model does not allow a closed-form solution; it must be approximated numerically. The numerical algorithm used here is based on a projection method with a collocation approach and is inspired by Fackler (2005) and Miranda and Glauber (1995). The results were obtained using MATLAB R2016b and solved using the solver for nonlinear rational expectations models RECS version 0.7 (Gouel, 2017). The numerical method is explained below, but for a complete picture see the program code (RECS code is publicly available and the code for this paper is available upon request). To simplify the exposition of the method, it is presented for a typical rational expectations model with informational subperiods, not for the specific equations of the model. Following Fackler (2005), rational expectations problems can be expressed using three groups of equations. For ease of exposition, we neglect below the year index, indexing only by

[11] This section draws in part on the appendix of Gouel and Jean (2015).

seasons. State variables s_i are updated through a transition equation:

$$s_i = g_{i-1}(s_{i-1}, x_{i-1}, e_i), \tag{A6}$$

where x_i are response variables and e_i are stochastic shocks. Response variables are defined by solving a system of complementarity equilibrium equations:

$$f_i(s_i, x_i, z_i) \leq 0, = 0 \text{ if } x_i > \underline{x}_i. \tag{A7}$$

Response variables can have lower bounds, $\underline{x}_i$. In cases where response variables have no bounds, equation (A7) simplifies to a traditional equation: $f_i(s_i, x_i, z_i) = 0$. z_i is a variable representing the expectations about next period and is defined by

$$z_i = \mathrm{E}_i \left[h_i \left(s_i, x_i, e_{i+1}, s_{i+1}, x_{i+1} \right) \right]. \tag{A8}$$

One way to solve this problem is to find a function that approximates well the behavior of response variables. We consider a cubic spline approximation of response variables,

$$x_i \approx \Phi_i \left(s_i, \theta_i \right), \tag{A9}$$

where θ_i are the parameters defining the spline approximation. To calculate this spline, we discretize the state space (using 150 nodes for availability and for expected production), and the spline has to hold exactly for all points of the grid.

The expectations operator in equation (A8) is approximated through 5-point Gaussian quadratures using functions from the CompEcon toolbox (Miranda and Fackler, 2002). The Gaussian quadrature defines a set of pairs $\left\{ e_i^l, w_i^l \right\}$ in which e_i^l represents a possible realization of shocks and w_i^l the associated probability. Using this discretization, and equations (A6) and (A8)–(A9), we can express the equilibrium equation (A7) as

$$f_i \left(s_i, x_i, \sum_l w_i^l h_i \left(s_i, x_i, e_{i+1}^l, g_i \left(s_i, x_i, e_{i+1}^l \right), \Phi_{i+1} \left(g_i \left(s_i, x_i, e_{i+1}^l \right), \theta_{i+1} \right) \right) \right) \leq 0, = 0 \text{ if } x_i > \underline{x}_i. \tag{A10}$$

For a given spline approximation, θ_{i+1}, and a given s_i, equation (A10) is a function of x_i and can be solved using a mixed complementarity solver.

Once all the above elements are defined, we can proceed to the algorithm, which runs as follows:

Step 1. Initialization step. Choose an initial spline approximation, θ_1^n with $n = 1$, based on a first-guess (for example, the steady-state values).

Step 2. Time iteration step for subperiods. For $i = I, \ldots, 1$ do

Step 2.1. Equation solving step. For each point of the grid of state variables, s_i^j, solve equation (A10) for

x_i^j using a mixed complementarity solver:

$$f_i\left(s_i^j, x_i^j, \sum_l w_i^l h_i\left(s_i^j, x_i^j, e_{i+1}^l, g_i\left(s_i^j, x_i^j, e_{i+1}^l\right), \Phi_{i+1}\left(g_i\left(s_i^j, x_i^j, e_{i+1}^l\right), \theta_{i+1}^n\right)\right)\right) \leq 0, = 0 \text{ if } x_i^j > \underline{x}_i. \tag{A11}$$

Step 2.2. Approximation step. Update the spline approximation using the new values of response variables, $x_i = \Phi_i\left(s_i, \theta_i^n\right)$.

Step 3. Terminal step. If $n = 1$ or $\|\theta^{n+1} - \theta^n\|_2 \geq 10^{-8}$ then increment n to $n+1$ and go to Step 2.

Once the rational expectations equilibrium is identified, the spline approximation of the decision rules can be used to simulate the model.

To ensure precise solutions, the simulations are done by solving exactly the equilibrium equation (A10) at each iteration using only the policy rules approximated by splines to approximate the expectations. Regarding the welfare results, since welfare terms are expressed as recursive equations such as equation (10), they are calculated by value function iterations. The value function iterations generate a function that represents welfare as a function of the state variables. This function is then applied to the simulated observations. The welfare in table 6 is the average of all welfare values in the first season.

Statistics on the asymptotic distribution are calculated over 1,000,000 observations from random outcomes of the stochastic variables, obtained by simulating 5,000 paths for 220 years and after discarding for each path the first 20 years as burn-in period. The aggregate production shocks are the same for all scenarios. This is ensured by first drawing shocks for the model with news, and then by multiplying them to obtain the aggregate shocks used in the scenarios without news or with all the news concentrated in the first period after planting.

References

Abbott, P., Boussios, D. and Lowenberg-DeBoer, J. (2016). *Valuing Public Information in Agricultural Commodity Markets: WASDE Corn Reports*. Conference paper, Agricultural and Applied Economics Association, 2016 Annual Meeting, July 31–August 2, Boston, Massachusetts.

Adjemian, M. K. (2012a). Emergence and impact of USDA's WASDE report. *Amber Waves*, June 05.

Adjemian, M. K. (2012b). Quantifying the WASDE announcement effect. *American Journal of Agricultural Economics*, 94(1), 238–256.

Adjemian, M. K., Johansson, R., McKenzie, A. and Thomsen, M. (2016). *The Value of Government Information in an Era of Declining Budgets*. Conference paper, Agricultural and Applied Economics Association, 2016 Annual Meeting, July 31–August 2, Boston, Massachusetts.

Beaudry, P. and Portier, F. (2014). News driven business cycles: Insight and challenges. *Journal of Economic Literature*, 52(4), 993–1074.

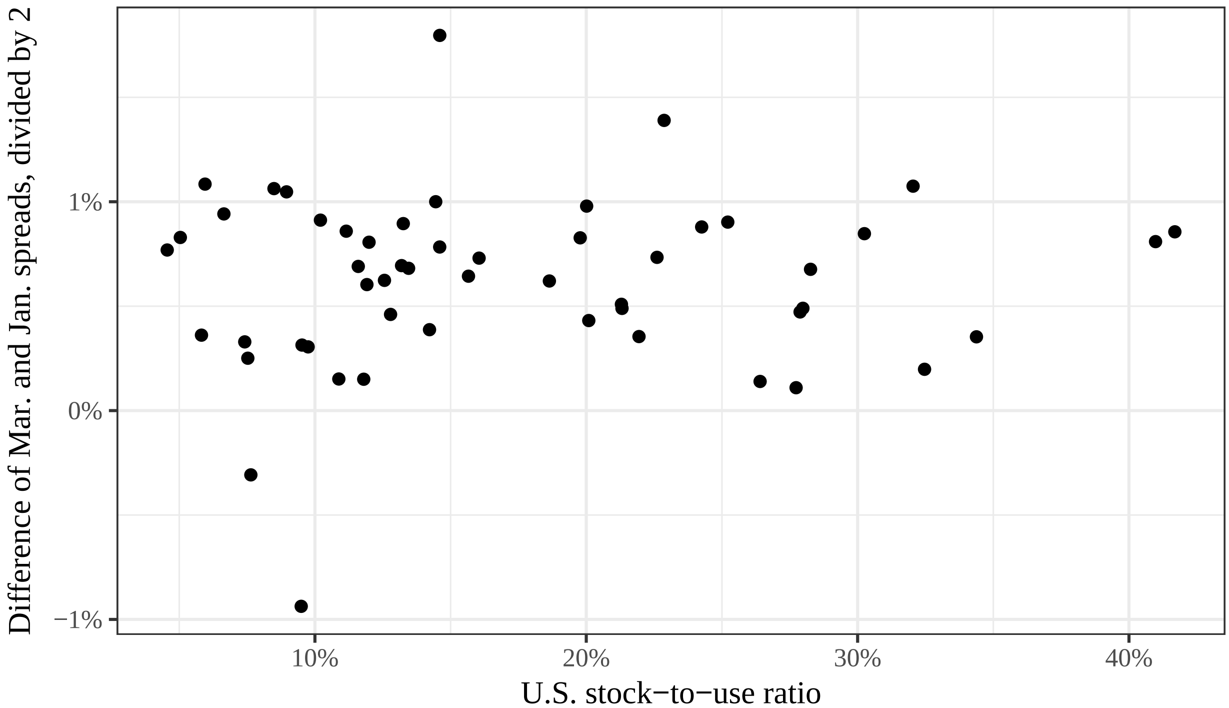

Figure A1: Differences of interest-adjusted March and January spreads, divided by 2, vs. inventory

Bontems, P. and Thomas, A. (2000). Information value and risk premium in agricultural production: The case of split nitrogen application for corn. *American Journal of Agricultural Economics*, 82(1), 59–70.

Bradford, D. F. and Kelejian, H. H. (1977). The value of information for crop forecasting in a market system: Some theoretical issues. *The Review of Economic Studies*, 44(3), 519–531.

Bradford, D. F. and Kelejian, H. H. (1978). The value of information for crop forecasting with bayesian speculators: Theory and empirical results. *The Bell Journal of Economics*, 9(1), 123–144.

Brennan, M. J. (1958). The supply of storage. *The American Economic Review*, 48(1), 50–72.

C-FARE (2016). *From Farm Income to Food Consumption: Valuing USDA Data Products*. Report, Council on Food, Agricultural & Resource Economics.

Cafiero, C., Bobenrieth, E. S. A., Bobenrieth, J. R. A. and Wright, B. D. (2011). The empirical relevance of the competitive storage model. *Journal of Econometrics*, 162(1), 44–54.

Chambers, M. J. and Bailey, R. E. (1996). A theory of commodity price fluctuations. *The Journal of Political Economy*, 104(5), 924–957.

Deaton, A. and Laroque, G. (1992). On the behaviour of commodity prices. *The Review of Economic Studies*, 59(1), 1–23.

Deaton, A. and Laroque, G. (1996). Competitive storage and commodity price dynamics. *The Journal of*

Political Economy, 104(5), 896–923.

Fackler, P. L. (2005). A MATLAB solver for nonlinear rational expectations models. *Computational Economics*, 26(2), 173–181.

FAO (2017). FAOSTAT database. Retrieved May 25, 2017 from http://www.fao.org/faostat/en.

Galvão, A. B. (2017). Data revisions and DSGE models. *Journal of Econometrics*, 196(1), 215–232.

Garcia, P., Irwin, S. H., Leuthold, R. M. and Yang, L. (1997). The value of public information in commodity futures markets. *Journal of Economic Behavior & Organization*, 32(4), 559–570.

Gouel, C. (2014). Food price volatility and domestic stabilization policies in developing countries. In J.-P. Chavas, D. Hummels and B. D. Wright (eds.) *The Economics of Food Price Volatility*, National Bureau of Economic Research conference report, chapter 7, (pp. 261–306). Chicago, IL: University of Chicago Press.

Gouel, C. (2017). RECS: MATLAB solver for rational expectations models with complementarity equations – version 0.7.

Gouel, C. and Jean, S. (2015). Optimal food price stabilization in a small open developing country. *World Bank Economic Review*, 29(1), 72–101.

Gouel, C. and Legrand, N. (2017). Estimating the competitive storage model with trending commodity prices. *Journal of Applied Econometrics*, 32(4), 744–763.

Guerra, E. A., Bobenrieth, E. S. A., Bobenrieth, J. R. A. and Cafiero, C. (2015). Empirical commodity storage model: The challenge of matching data and theory. *European Review of Agricultural Economics*, 42(4), 607–623.

Haile, M. G., Kalkuhl, M. and von Braun, J. (2016). Worldwide acreage and yield response to international price change and volatility: A dynamic panel data analysis for wheat, rice, corn, and soybeans. *American Journal of Agricultural Economics*, 98(1), 172–190.

Harrell, F. E. (2001). *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. Springer Series in Statistics, New York: Springer.

Hayami, Y. and Peterson, W. (1972). Social returns to public information services: Statistical reporting of U.S. farm commodities. *The American Economic Review*, 62(1/2), 119–130.

Hendricks, N. P., Janzen, J. P. and Smith, A. (2015). Futures prices in supply analysis: Are instrumental variables necessary? *American Journal of Agricultural Economics*, 97(1), 22–39.

Huang, H.-H. and Moore, M. R. (forthcoming). Farming under weather risk: Adaptation, moral hazard, and selection on moral hazard. In W. Schlenker (ed.) *Understanding Productivity Growth in Agriculture*, National Bureau of Economic Research conference report. Chicago, IL: University of Chicago Press.

IMF (2014). *World Economic Outlook: Recovery Strengthens, Remains Uneven*. Washington, DC: International Monetary Fund.

Kaldor, N. (1939). Speculation and economic stability. *The Review of Economic Studies*, 7(1), 1–27.

Karali, B. and Thurman, W. N. (2009). Announcement effects and the theory of storage: an empirical study of lumber futures. *Agricultural Economics*, 40(4), 421–436.

Lechthaler, F. and Vinogradova, A. (2017). The climate challenge for agriculture and the value of climate services: Application to coffee-farming in Peru. *European Economic Review*, 94, 45–70.

Leduc, S., Moran, K. and Vigfusson, R. J. (2016). *Learning in the Oil Futures Markets: Evidence and Macroeconomic Implications*. International Finance Discussion Paper 1179, FRB.

Lowry, M., Glauber, J. W., Miranda, M. J. and Helmberger, P. G. (1987). Pricing and storage of field crops: A quarterly model applied to soybeans. *American Journal of Agricultural Economics*, 69(4), 740–749.

Lybbert, T. J., Smith, A. and Sumner, D. A. (2014). Weather shocks and inter-hemispheric supply responses: Implications for climate change effects on global food markets. *Climate Change Economics*, 5(4), 1450010.

Merener, N. (2015). Globally distributed production and the pricing of CME commodity futures. *Journal of Futures Markets*, 35(1), 1–30.

Miranda, M. J. and Fackler, P. L. (2002). *Applied Computational Economics and Finance*. Cambridge, US: MIT Press.

Miranda, M. J. and Glauber, J. W. (1995). Solving stochastic models of competitive storage and trade by Chebychev collocation methods. *Agricultural and Resource Economics Review*, 24(1), 70–77.

Ng, S. and Ruge-Murcia, F. J. (2000). Explaining the persistence of commodity prices. *Computational Economics*, 16(1–2), 149–171.

Osborne, T. (2004). Market news in commodity price theory: Application to the Ethiopian grain market. *The Review of Economic Studies*, 71(1), 133–164.

Paul, A. B. (1970). The pricing of binspace: A contribution to the theory of storage. *American Journal of Agricultural Economics*, 52(1), 1–12.

Peterson, H. H. and Tomek, W. G. (2005). How much of commodity price behavior can a rational expectations storage model explain? *Agricultural Economics*, 33(3), 289–303.

Roberts, M. J. and Schlenker, W. (2013). Identifying supply and demand elasticities of agricultural commodities: Implications for the US ethanol mandate. *The American Economic Review*, 103(6), 2265–2295.

Sacks, W. J., Deryng, D., Foley, J. A. and Ramankutty, N. (2010). Crop planting dates: an analysis of global patterns. *Global Ecology and Biogeography*, 19(5), 607–620.

Steinwender, C. (forthcoming). Information frictions and the law of one price: "when the States and the Kingdom became united". *The American Economic Review*.

USDA (2010). *Field Crops: Usual Planting and Harvest Dates*. Agricultural Handbook 628, United States Department of Agriculture, National Agricultural Statistics Service.

USDA (2017a). Production, Supply and Distribution (PSD) data. Released on July 12, 2017 and retrieved from `https://apps.fas.usda.gov/psdonline/app/index.html`.

USDA (2017b). World Agricultural Supply and Demand Estimates Report (WASDE). Retrieved from https://www.usda.gov/oce/commodity/wasde/.

Williams, J. C. and Wright, B. D. (1991). *Storage and Commodity Markets*. New York: Cambridge University Press.

Winne, J. D. and Peersman, G. (2016). Macroeconomic effects of disruptions in global food commodity markets: Evidence for the United States. *Brookings Papers on Economic Activity*, 47(2), 183–286.

Working, H. (1949). The theory of price of storage. *The American Economic Review*, 39(6), 1254–1262.