



HAL
open science

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

Christophe C. Gouel

► **To cite this version:**

Christophe C. Gouel. The Value of Public Information in Storable Commodity Markets: Application to the Soybean Market. *American Journal of Agricultural Economics*, 2020, 102 (3), pp.846-865. 10.1002/ajae.12013 . hal-02622352v2

HAL Id: hal-02622352

<https://hal.inrae.fr/hal-02622352v2>

Submitted on 18 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - ShareAlike | 4.0 International License

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

Christophe Gouel^{1,2}

¹Économie Publique, AgroParisTech, INRA, Université Paris-Saclay, 78850, Thiverval-Grignon, France

²CREST, École Polytechnique, 91120, Palaiseau, France

Abstract

This article provides a framework to estimate the potential effects and benefits of the provision of market information in storable commodity markets. This framework is applied to the case of production forecasts for 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. Shocks that occur between planting and harvesting affect the size of the potential harvest. Estimates of the size of these shocks are reported publicly, and affect the market equilibrium through adjustments to stock levels. By varying counterfactually the observability of seasonal shocks, we can estimate the efficiency gains related to the availability of advance information. They are equivalent to 2% of storage costs; the reduction of stock levels being the main channel explaining the welfare gains. The presence of advance information has a limited effect on inter-annual price volatility but redistributes price volatility during the season, increasing it just before harvest when almost all news has been received and stocks are tight, and decreasing it after harvest. 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.

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

JEL classification: D84, Q02, Q11, Q13.

*Christophe Gouel is a senior research fellow at INRA, Économie Publique. The author is grateful to the editor Timothy Richards and two anonymous referees, as well as Tamara Ben Ari, Benjamin Carton, Joe Glauber, Dirk Krueger, Nicolas Legrand, Marc Nerlove, Will Martin, and Steve Salant for helpful comments. Special thanks to Scott Irwin for his insightful comments and for providing data on private forecasts. This work was supported by the INRA-Cirad metaprogramme GloFoodS and has been undertaken as a part of the CGIAR Research Programs on Policies, Institutions, and Markets (PIM), which is led by the International Food Policy Research Institute (IFPRI) and funded by the CGIAR Fund Donors. Christopher Marciniak and Maimouna Yokessa provided excellent research assistance. Correspondence to be sent to christophe.gouel@inra.fr.

This is the peer reviewed version of the following article: “Gouel, C. (2020). The Value of Public Information in Storable Commodity Markets: Application to the Soybean Market. *American Journal of Agricultural Economics*, 102(3) 846–865”, which has been published in final form at <https://doi.org/10.1002/ajae.12013>. This article may be used for non-commercial purposes in accordance with [Wiley Terms and Conditions for Use of Self-Archived Versions](#).

Accurate and timely information is key to the good functioning of financial markets including commodity markets. What is specific to commodity markets is the extensive public involvement in the provision of market information. For example, national and international agencies regularly report stock levels and, in the case of agricultural commodities, forecast the coming harvests. 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 Agriculture Ministers as a response to the global food price hikes that occurred in 2007/08 and 2010. The important role of public information has been acknowledged in the literature (C-FARE, 2016) but evaluations of its counterfactual effects and benefits are rare despite the costs to public agencies of collecting data through surveys which is very high.

The objective of this article is to develop 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 efficient market allocation, by leveling the playing field for market participants with unequal access to private information, and by providing a basis for research on market behavior. Here, we focus on the allocation of resources over time and how it is affected by the timing of information. We build a rational expectations storage model that accounts for both inter-annual and intra-annual market dynamics, and in which shocks between planting and harvesting affect the size of the potential harvest.

We apply the model to the case of soybean because its production is very concentrated—more than 80% of world production comes from Argentina, Brazil, and the United States (U.S.)—and its market is well integrated internationally (e.g., Merener, 2015, shows that soybean futures react similarly to weather events in Argentina, Brazil, and the U.S.), two features that simplify the modeling. To make the model computationally tractable, we define a simplified crop calendar corresponding to the most representative planting and harvesting dates and use monthly changes in U.S. Department of Agriculture (USDA) production forecasts to calibrate the size of the news shocks. Varying the size of the observable seasonal shocks counterfactually allows an assessment of the benefits of providing public information about potential shocks.

The effects of news shocks are analyzed using a rational expectations storage model extended to account simultaneously for inter-annual and intra-annual market dynamics. Adopting a rational expectations framework is a natural starting point as it implies that agents optimally process the information provided by news shocks.¹ Our model builds on the previous work with rational expectations storage models featuring intra-annual dynamics, in particular the three papers in this literature that include news shocks (Ng and Ruge-Murcia, 2000; Osborne, 2004; Peterson and Tomek, 2005). Here, the word “news” takes the meaning it has been given in the macroeconomic literature (Beaudry and Portier, 2014) of observation of exogenous shocks with effects on fundamentals in the future.² In order to capture the rich effects of news shocks on price dynamics, we build a monthly model, where each month is characterized by a different state of the world and different action sets. There are planting months, harvesting months, and months where agents receive news about the coming harvests.

¹The objectives of the recent work by Abbott et al. (2016) are similar to ours but without rational expectations, which makes it unclear how agents process the new information.

²The early work by Lowry et al. (1987), Williams and Wright (1991, Ch. 8), and Chambers and Bailey (1996) features intra-annual storage dynamics without news shocks.

Except for the monthly time-frame and the representation of supply as elastic, our model is very close to Osborne (2004). Osborne (2004) studies how precipitation observed in advance of harvest affects the market equilibrium in a storage model for the Ethiopian grain market. Our monthly model (compared to Osborne's quarterly model) allows us to show that news shocks affect price volatility in subtle ways: they slightly reduce the overall price volatility but their main effect is to 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, we do not have to build the news shocks from meteorological information; our reliance on USDA production forecasts means we have direct observations of news shocks expressed in quantities which allows better identification of the information available to agents.

To calibrate the size of the news shocks, we use USDA production forecasts because their monthly releases are watched closely by market observers, and are considered a benchmark. USDA releases reports on crop supply and demand on around the 10th of every month (see figure A4 in the appendix for an example of the monthly balance sheets). Our study uses only production forecasts because after planting, changes to production forecasts can reasonably be considered exogenous and so can be seen as news shocks. In 1980, the most important foreign markets for each commodity began to be included in the balance sheets, which provides us with changes in production forecasts for Argentina and Brazil. Other institutions than USDA—private firms and international organizations such as the International Grains Council, and since 2011, AMIS—also produce crop forecasts. So, assuming a counterfactual without any advance information would overestimate the effects of USDA reports. We use a series of private forecasts about the U.S. harvest to decompose the news shocks between components that are observed by all forecasters and components that are observed only by USDA. This decomposition allows us to differentiate the total effect of the existence of news shocks from the increased precision allowed by the USDA forecasts.

To make the storage model tractable, we make two key simplifications. First, we assume an integrated world market without trade costs. This assumption decreases the potential benefits from news shocks by neglecting the possibility to adjust costly trade flows to the arrival of information (as in Steinwender, 2018), but it also reduces the model's size by not having to track separate state variables in different countries. Second, we represent only two producing regions, the U.S. and an aggregate of Argentina and Brazil, for which we have more than 30 years of production forecasts. The model does not include smaller producers, the largest of which being China with 4% of world production. In the current setting, including China would change little to the results. Given that for soybean China has a crop calendar similar to the U.S., and that the model assumes an integrated world market, including news shocks related to China could be done simply by averaging the U.S. and Chinese news shocks together. However, since China production is one-eighth of the U.S. one, the news shocks related to the U.S. production would dominate the results.³ News shocks regarding the Chinese market could matter much more if considering other dimensions of the balance sheet (e.g., domestic use, import, ending stocks), but these dimensions mix endogenous market responses and news shocks, making them much more difficult to account for.

The market effect of USDA production forecasts is the object of a large empirical literature, inspired by market efficiency studies in finance and assessing whether and how markets react to public announcements

³Even more so if accounting for the fact that most of China production is non-GMO food-grade soybean that does not directly compete on the main soybean market.

(e.g., [Garcia et al., 1997](#)). This literature investigates the value of USDA forecasts through their capacity to affect the market. USDA market outlooks are not the only source of market information. The existence of other private sources of information raises questions about the economic function of public information. One way to assess its role is to quantify market reactions to the release of USDA reports. [Adjemian \(2012\)](#) and [Adjemian et al. \(2018\)](#) show that agricultural markets do react to the information contained in USDA reports, and that their reactions are consistent with storage theory.⁴ However, although this literature quantifies how markets react to the release of public information, it is silent on the public benefits of this information, the object of the present study. Our article complements this empirical literature by providing a theoretical framework to interpret the effects of these production forecasts and by allowing the possibility to analyze counterfactual situations.

Given our model's structure, the gains to expect from advance information will come from better allocation of stocks through time. This was also the main mechanism studied in the early approaches proposed in [Hayami and Peterson \(1972\)](#) and [Bradford and Kelejian \(1977, 1978\)](#). These works developed analytical frameworks based on intra-annual storage models to assess the welfare costs of measurement errors in public information. Our approach differs in that we include the problem of seasonal news shocks in a modern rational expectations storage model such that seasonal information matters for both seasonal market dynamics and inter-annual market dynamics, because advance information about harvests allows smoother inter-annual stock adjustments.

Adjustment of inter-annual stock is likely to be one of the main channels through which USDA reports affect resource allocation, but it is not the only possible one. For example, [Hayami and Peterson \(1972\)](#) also consider the potential gains coming from adjustments of production to news shocks. In our model, we assume that supply is elastic at planting but that after that time it is not possible for production to react to shocks (see [Bontems and Thomas, 2000](#), and [Lechthaler and Vinogradova, 2017](#), for work on this issue). Since soybean, as a leguminous plant, does not require nitrogen fertilization during the growing season and is mostly non-irrigated, the assumption of inelastic supply after planting is reasonable for this crop. This assumption allows a clean identification of the exogenous shocks from the available crop production forecasts. For other commodities, for which some inputs can be adjusted after production levels have been planned, this assumption can be easily lifted. More generally, the news shocks could also have effects on resources allocation outside the sector under study. For example, by affecting livestock feed demand or planting of substitute crops. So, although capturing one of the main adjustment channel, our results likely underestimate the welfare impacts of these news shocks.

A seasonal storage model

We consider the soybean world market represented by 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 $r \in \{\text{US}, \text{LAC}\}$,

⁴[Karali and Thurman \(2009\)](#) show that the lumber futures market also reacts to announcements in a way that is consistent with the storage model.

corresponding to the U.S., and Argentina and Brazil combined. The two regions differ in 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.

Soybean generally takes around five months to grow from sowing to harvest. However, because of the variety of local climates, sowing to harvesting in the U.S. extends from mid-April to late December. The pattern is similar in Argentina and Brazil. To simplify the model, we assume that the U.S. production takes place over the five most active months from May/June to October/November, and the Argentina-Brazil production takes place between October/November and March/April. This simplified crop calendar means that the crop cycles do not overlap between regions: when planting occurs in one country harvest has been completed in the other. Since producers plant after the revelation of the size of the other hemisphere harvest, their price incentives are affected by this information through the effect of stock level adjustments on expected price, and their production plan are adjusted accordingly.

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: $D(P_{i,t})$.

Storers

There is a single representative speculative storer that 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 to 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

$$(1) \quad \beta E_{i,t} P_{i+1,t} - P_{i,t} - k\bar{P} \leq 0, = 0 \text{ if } S_{i,t} > 0,$$

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

Because of the convention that $\{I + 1, t\} = \{1, t + 1\}$, equation (1) applies equally to storers' behavior between seasons and between years.

Producers

Production in each region is undertaken by a representative competitive producer that takes the planting decision before knowing the selling price and the yield. Production technology is at constant return to scale but requires a scarce and heterogeneous input, land, so that increasing production requires increasing the

use of less fertile fields which increases in turn the marginal production cost. For simplicity, we collapse the producer and landowner, and represent the constraint placed by land supply by a convex production cost function. The producer in region r plants in season i^r in year t in the expectation of harvesting the quantity $Q_{i^r,t}^r$ five months later. However, the production is affected by a multiplicative random shock $\epsilon_{i^r+5,t}^r$ which follows a distribution with unitary mean that is described later. We assume that after planting the producer cannot adjust its production level. The producer chooses its production level by solving the following maximization of expected profit:

$$(2) \quad \max_{Q_{i^r,t}^r} \beta^5 E_{i^r,t} \left(P_{i^r+5,t} \epsilon_{i^r+5,t}^r Q_{i^r,t}^r \right) - \Psi^r \left(Q_{i^r,t}^r \right),$$

where $\Psi^r \left(Q_{i^r,t}^r \right)$ is the convex cost of planning the production $Q_{i^r,t}^r$ and $\epsilon_{i^r+5,t}^r Q_{i^r,t}^r$ is the realized production level. Profit maximization gives the following intertemporal equation

$$(3) \quad \beta^5 E_{i^r,t} \left(P_{i^r+5,t} \epsilon_{i^r+5,t}^r \right) = \Psi^{r'} \left(Q_{i^r,t}^r \right),$$

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

News shocks

In contrast to 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 such that they provide advance information on the size of the coming harvest. They are news shocks in the sense that they do not directly affect current quantities but they do affect expectations about realizations of future quantities. Nonetheless, they affect the market equilibrium through immediate adjustments to stock levels. A higher than average news shock will increase expected output, decreasing in turn expected price at harvest and incentives to store. To maintain the non-arbitrage condition (1), storers will sell more to consumers driving the current price lower. The converse applies with a lower than average news shocks.

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

$$(4) \quad \epsilon_{i^r+5,t}^r = \exp \left(\sum_{i=1}^5 \eta_{i^r+i,t}^r \right),$$

where the $\eta_{i,t}^r$ are i.i.d. and follow normal distributions with mean μ_i^r and standard deviations σ_i^r such that $\mu_i^r + (\sigma_i^r)^2/2 = 0$ which ensures that seasonal shocks have unitary mean. Under these assumptions, $\epsilon_{i^r+5,t}^r$

⁵We adopt a representation with 5 shocks because, for the U.S., each shock will correspond to changes in production forecast between two WASDE reports. In addition, depending on the varieties, the time to maturity of soybean is 90 to 150 days, so 5 months would cover the entire growing season of a long-to-mature soybean variety.

follow a lognormal distribution with parameters $\sum_{i=1}^5 \mu_{i^r+i}^r$ and $\sqrt{\sum_{i=1}^5 (\sigma_{i^r+i}^r)^2}$. Given that $\mu_i^r + (\sigma_i^r)^2 / 2 = 0$, this implies that the expected mean of $\epsilon_{i^r+5,t}^r$ is 1.

There is no need for an explicit representation of shocks that are unobservable. In assuming that production is perfectly observed after harvest, the implication is that the unobservable part of the seasonal shocks is shifted to the harvest season, so in $\eta_{i^r+5,t}^r$, when it becomes observable.

The assumption of i.i.d. news shocks implies weakly efficient monthly production forecasts (Nordhaus, 1987). Monthly production shocks could be correlated but efficient forecasters should account for this correlation in forming their forecasts resulting in independent forecast revisions. However, there is some evidence that revisions to USDA crop production forecasts are smoothed (Isengildina et al., 2006), but the correlation identified for soybean being small it is neglected here.

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

$$(5) \quad \hat{Q}_{i,t}^r = E_{i,t} \left[Q_{i^r,t}^r \exp \left(\sum_{j=i^r+1}^{i^r+5} \eta_{j,t}^r \right) \right] = Q_{i^r,t}^r \exp \left(\sum_{j=i^r+1}^i \eta_{j,t}^r \right).$$

Market equilibrium

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

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

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:

$$(7) \quad A_{i,t} = \begin{cases} S_{i-1,t} + \epsilon_{i,t}^r Q_{i-5,t}^r & \text{if } i = i^r + 5, \\ S_{i-1,t} & \text{if } i \neq i^r + 5. \end{cases}$$

In seasons after planting but before harvesting, defined by the set $\mathcal{I} = \{i^r + j | j = \{1, \dots, 4\} \text{ and } r = \{\text{US, LAC}\}\}$, the expected production, $\hat{Q}_{i,t}^r$, is also a state variable since it will affect resource allocation via storage. Past storage and expected production play similar roles in characterizing the market supply but can be summed only at harvest time. Until then, expected production is still uncertain and so is not equivalent to stocks. In addition, even if expected production were 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.

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

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

Definition. A recursive equilibrium is a set of functions $\mathcal{S}_i (s_{i,t})$, $\mathcal{Q}_{i^r}^r (s_{i^r,t})$, and $\mathcal{P}_i (s_{i,t})$, 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.

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. However, 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:⁶

$$(8) \quad w_{i,t} = \int_{P_{i,t}}^{P_{\max}} D(p) \, dp + P_{i,t} S_{i,t-1} - (k\bar{P} + P_{i,t}) S_{i,t} + \begin{cases} P_{i,t} \epsilon_{i,t}^r Q_{i-5,t}^r & \text{if } i = i^r + 5, \\ -\Psi^r(Q_{i,t}^r) & \text{if } i = i^r, \\ 0 & \text{if } i \notin \{i^r, i^r + 5\}, \end{cases}$$

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

$$(9) \quad w_{i,t} = \underbrace{\int_{P_{i,t}}^{P_{\max}} D(p) \, dp + P_{i,t} (A_{i,t} - S_{i,t}) - k\bar{P} S_{i,t}}_{\text{Consumers' efficiency gains}} - \mathbf{1}_{i=i^r} \Psi^r(Q_{i,t}^r),$$

where $\mathbf{1}_{(\cdot)}$ is the indicator function. 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 help understanding the source of welfare changes.

From instantaneous welfare, we can calculate the intertemporal welfare, $\tilde{W}_{i,t}$, by

$$(10) \quad \tilde{W}_{i,t} = w_{i,t} + \beta E_{i,t} \tilde{W}_{i+1,t}.$$

To obtain a value of welfare that is directly comparable to the value of monthly consumption, we introduce intertemporal welfare normalized to a monthly value as $W_{i,t} = (1 - \beta) \tilde{W}_{i,t}$, which using equation (10) gives

$$(11) \quad W_{i,t} = (1 - \beta) w_{i,t} + \beta E_{i,t} W_{i+1,t}.$$

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

⁶When producing soybean, production costs are not all incurred at planting; harvesting is also a significant source of costs. As long as production is inelastic after planting, the exact timing of costs does not matter for the results and, for simplicity, we assume they are all incurred at planting when the production decision is taken.

consequences for the results, we choose the first season to calculate the welfare results. Welfare is evaluated using value function iterations once the solution of the storage model has been found (see online appendix for details on the numerical methods).

Taking the model to the data

To take the model to the data, we assume functional forms to be isoelastic, with demand given by

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

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

$$(13) \quad \Psi^r(Q^r) = \frac{\beta^5 \bar{P}}{(\theta^r \bar{D})^{1/\alpha^Q}} \frac{(Q^r)^{1+1/\alpha^Q}}{1+1/\alpha^Q},$$

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 to be the same in all regions. With this expression of production costs, equation (3) simplifies to

$$(14) \quad Q_{i^r,t}^r = \theta^r \bar{D} \left[E_{i^r,t} \left(\frac{P_{i^r+5,t}}{\bar{P}} \epsilon_{i^r+5,t}^r \right) \right]^{\alpha^Q}.$$

\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 into one period, and production and stockpiling decisions taken inter-annually (see the online appendix). It is easier to work with these two annual steady-state values than with the 24 seasonal values of the monthly model.

Using these functional forms where variables are expressed in deviation from their respective steady-state annual value, the variables in levels can be substituted in all the equations by the corresponding ratio to the annual steady-state levels (see equations (A6)–(A11) in online appendix for a proof). Since all variables, including welfare, can be expressed unit-free, it means that \bar{P} and \bar{D} serve only to define the level of prices and quantities without affecting any of the other results and we can normalize them to any values. Without loss of generality, we take $\bar{P} = 1$ and $\bar{D} = 12$. 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 U.S., θ^{US} , is 42%, its average 2007–2016 value (considering that Argentina, Brazil, and the U.S. are the only producers).

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})	0.56	
θ^{US}	Share of US production in total production (%)	42	
α^D	Demand elasticity	-0.2	
α^Q	Supply elasticity	0.4	
		US	LAC
$\sigma_{i^r+1}^r$	Standard deviation of the 1 st news shocks (%)	3.65	2.96
$\sigma_{i^r+2}^r$	Standard deviation of the 2 nd news shocks (%)	4.31	3.21
$\sigma_{i^r+3}^r$	Standard deviation of the 3 rd news shocks (%)	3.63	2.29
$\sigma_{i^r+4}^r$	Standard deviation of the 4 th news shocks (%)	3.33	2.56
$\sigma_{i^r+5}^r$	Standard deviation of the 5 th news shocks (%)	1.73	2.43
$\sqrt{\sum_{i=1}^5 (\sigma_{i^r+i}^r)^2}$	Standard deviation of the total production shock (%)	7.69	6.07

Behavioral parameters

This storage model is too complex and takes too long to solve to be amenable to a fully structural estimation with cross-equation restrictions. Instead, we estimate supply elasticity with an approach that follows [Roberts and Schlenker \(2013\)](#). The estimation strategy and the results are available in the online appendix.⁷ The estimation is based on equation (14), multiplied by the production shock, and expressed in log. The conditional expectation of the harvest-time price multiplied by the production shock, which is not observable, is replaced by the expected price, which empirically is assumed to correspond to the futures price at planting time for delivery at harvest. Adding a region-specific cubic-spline trend leads to the main equation to estimate supply elasticity. Following [Roberts and Schlenker \(2013\)](#) and to account for the fact that futures prices can reflect anticipated shocks unobserved to the econometrician, they are instrumented by lagged production shocks and lagged prices. The preferred 2SLS estimates implies a supply elasticity of 0.4.

We have not been able to estimate soybean demand elasticity using [Roberts and Schlenker's](#) method. They obtain a statistically significant demand elasticity estimating a four-crop system, only by imposing symmetry and by using USDA rather than FAOSTAT data (table 8). They obtain an elasticity of -0.236. We are unaware of any other recent source of credibly estimated demand elasticity for soybean, so we follow [Roberts and Schlenker](#) and use a demand elasticity of -0.2.

Storage cost is recovered from the maximum spreads between two futures contract: the nearby contract and the next-to-expire contract (see online appendix). We settle on a value of 5 cents per bushel per month which using an average price of \$8.98 per bushel implies $k = 0.0056$.

The influence on the results of each of the three behavioral parameter is evaluated in a sensitivity analysis.

⁷Data and programs to replicate the estimations and the simulations are available at DOI: [10.15454/YWAOKM](https://doi.org/10.15454/YWAOKM).

Seasonal shocks

To calibrate the size of the seasonal shocks, we consider that there are two sources of information. One is the monthly changes in production projections from the World Agricultural Supply and Demand Estimates (WASDE) reports (USDA, 2017b). The other is the consensus forecast from private analysts.

Public forecasts

The WASDE reports are balance sheets consolidating the information available in different USDA services. For the forecasts regarding U.S. production, the original information comes from the National Agricultural Statistics Service (NASS). The forecasting cycle for soybean works as follows (Good and Irwin, 2006). It begins with the Prospective Plantings report that is released at the end of March and that contains the farmers' planting intentions. Planted acreages are then updated at the end of June in the Acreage report. In the following months, estimates of acreages will be adjusted only slightly, and the focus shifts to yields forecasts. Around the 10th of every month, NASS releases a Crop Production report, the main elements of which can be found in the WASDE report released at the same moment and used in this article. The August report contains the first non-trend yield forecast, and these forecasts are updated in the following months. Before the August report, the U.S. production forecast is the product of the acreage estimates by the trend yields. The production forecasts for foreign countries have a less rigid schedule and are based on information from USDA Foreign Agricultural Service attaché reports, foreign official sources, and satellite imagery, among other sources.

To make the link between the model and the data, we note that using equation (5), we obtain $\exp(\eta_{i,t}^r) = \hat{Q}_{i,t}^r / \hat{Q}_{i-1,t}^r$: the seasonal shocks can be identified by calculating the ratio between two consecutive projections. Given the assumption of shocks with unitary mean, we have $\text{var}(\hat{Q}_{i,t}^r / \hat{Q}_{i-1,t}^r) = \exp((\sigma_i^r)^2) - 1 \approx (\sigma_i^r)^2$. One benefit of assuming multiplicative shocks is that we do not need to detrend the data to compare shocks across years regardless of the trending production level.⁸ Based on these seasonal shocks, table 2 reports the mean and standard deviation of the monthly adjustments to production projections. We cannot reject the null hypothesis that the mean is equal to 1, as is assumed in the model, for any of the months considered. Therefore, the standard deviation can be interpreted as a coefficient of variation and is multiplied by 100 to ease interpretation.

For the U.S. market, projections for the new harvest begin in the WASDE reports with planting in the reports for May (in earlier reports, projections for the new crop year started in June or July) based on the combination of information on planting intentions and trend yields. There are no changes in projections between May and June, as acreage and yield estimates are not updated. The normal planting and harvesting periods for U.S. soybean are as follows (USDA, 2010). Planting of soybean begins in late April/early May and ends in early July with second half of May being the most active period. U.S. soybean harvest begins in September and is mostly completed by end November. This is consistent with the WASDE reports. Most production changes occur between June and November, with the largest changes observed in the August report which is the first to include survey-based yield forecasts and not just trend yields. Between June and July,

⁸The presence of a trend in production, and even more so in prices, alters the storage incentives. However, while this is of importance in estimating the storage model (Bobenrieth et al., 2017), accounting for such trends are unlikely to affect the results in this paper.

Table 2. Mean and Standard Deviation of Monthly Adjustments to Production Projections

Month	U.S. (1973–2016) ^a				Argentina and Brazil (1984–2016) ^a				
	# Obs.	Mean	SD (%)	# Obs. ^d	Argentina		Brazil		LAC ^b
					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). If there is more than one production forecast per month, as in the 1970s, we select the release date that is closest to the 10th of the month.

Note: Dashed segments delineate the months used for calibration. ^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. Since they would not affect any of the results, they were not collected manually.

yield estimates do not change and all changes in production are explained by changes in planted and harvested acreages. After July, most of the adjustments come from changes in yields, acreage estimates being fairly stable. No adjustments are made to production projections between November and December; some small adjustments are made between December and March based on data revisions. Further revisions can occur several months later as 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. soybean is harvested corresponds to the period October 15 to November 15, denoted October/November. Therefore, in the model we represent the U.S. soybean crop cycle as starting with the planting decision in May/June and ending with harvesting in October/November. There are five months following planting each with a seasonal shock. The standard deviations of the seasonal shocks, σ_i^{US} ,

are simply the standard deviations of the adjustments to production between July and November (table 2).

Using these five standard deviations for seasonal shocks implies a standard deviation for the equivalent aggregate shock of 7.7%. This number can be compared to a between-year volatility in production of 14.8% (standard deviation of the first-order difference of the logarithm of final U.S. production values as reported in USDA Production, Supply and Distribution database): the combined seasonal shocks appear to account for a sizable share of this volatility. The remaining production volatility is attributable to shocks occurring before planting, the endogenous reactions of farmers to market incentives, unobservable seasonal shocks that are revealed in the data revisions made after harvest, and other types of shocks such as input price shocks.

For Argentina-Brazil, identification of the seasonal shock structure is more challenging. Taken together, these countries present much greater agro-climatic diversity than the U.S., so the planting and harvesting of the main soybean crop is spread over more time. In addition, because of the tropical climate in these countries, double-cropping systems are possible, spreading soybean production even further over time. The WASDE reports began separate coverage of Argentina and Brazil only in January 1985. Therefore, we have fewer observations than for the U.S. The production projections for these countries present different patterns from those for the U.S. They start with the U.S. projections, in May in the most recent reports, so several months before planting in the Southern hemisphere. The initial projections are estimates based on likely production level trends by USDA oilseeds experts. These early projections can be adjusted before the planting seasons, based on additional information. Another feature that differs from U.S. projections is that the projections for Argentina and Brazil do not seem to converge smoothly to some final estimates, even in the case of August, some months after the end of the main harvest. Crop adjustments are much larger in Argentina than in Brazil or the U.S., but because it is a smaller producer than Brazil and their production shocks are not correlated, the volatility of the Argentina-Brazil region is often smaller than the volatility of the countries considered separately.

In order to obtain a series of five seasonal shocks for this region, we have to make some assumptions. Based on the AMIS crop calendar,⁹ most planting in Brazil happens in October and November, and the harvest takes place in March and April. In Argentina, planting is during November and December, and harvesting takes place in April and May. For this region, the model assumes that planting takes place in October/November and harvesting takes place in March/April. We consider that the main news shocks are the five shocks with the largest standard deviations between January and May. However, were we to use the values for these months in the last column of table 2, this would result in a standard deviation of the total production shock of 4.74%, well below the 11.67% that is observed inter-annually. Consequently, we rescale seasonal shocks so that the total production shock represents 52% of the inter-annual volatility, which is observed for the U.S. The resulting standard deviations are presented in table 1. They show a more regularly spread variability than in the case of the U.S.

Private forecasts

In addition to USDA forecasts, there are also private analysts who produce their own harvest forecasts and release them to their clients a few days before the USDA reports. To avoid overestimating the gains from the

⁹<http://www.amis-outlook.org/amis-about/calendars/en/>, retrieved March 27, 2017.

provision of public information, it is important to be able to account for the information that would come from these other sources in the absence of USDA reports. But were the private forecasts really private, they would not affect our assessment of the benefits of public information. Actually, they do not stay private for long. After their initial private release based on which traders adjust their positions, private forecasts are often made public in specialized professional news outlets (e.g., Bloomberg and Reuters) through surveys of the traders or directly from the firms establishing the forecasts.

We use a database of private forecasts about the U.S. harvest extending from 1983 to 2017 for the July forecasts and from 1970 to 2017 for the August, September, October, and November forecasts.¹⁰ There is no equivalent database for the Argentina and Brazil crop forecasts, but their production is also monitored by some market analysts. What we have called the July forecast is different from the forecast in the later months, because it is concerned only with forecasting the acreages and is released in the NASS Acreage report at the end of June before being incorporated in the July WASDE report. For 1983–2015, the July forecasts are an average of analyst forecasts collected by Bill Tierney. After 2015, it comes from Reuters. For 1970–2000, the August, September, October, and November forecasts are obtained as a simple average of the production forecasts by Conrad Leslie and Informa Economics (formerly Sparks Companies), selected because they were considered to be the most influential, they were widely reported and they are available over an extended period of time. For 2001–5, the forecasts are a simple average of Informa Economics estimate and the average analyst estimate reported by the Dow Jones Newswire survey. For 2006–12, the forecasts are the Dow Jones survey average. For 2013–5, the forecasts are the Bloomberg survey average. For 2016–7, the forecasts are from Reuters. See Egelkraut et al. (2003) and Good and Irwin (2006) for details about these private forecasts and their performance.

To identify the private shocks, we assume that the monthly shocks to production are the sum of two elements: a part observed by both private forecasting firms and USDA, denoted $\tilde{\eta}_{i,t}^r$, and a part observed by USDA only, denoted $\eta_{i,t}^r$, so that the total shock is the sum of the two, $\eta_{i,t}^r = \tilde{\eta}_{i,t}^r + \eta_{i,t}^r$. These two shocks follow normal distributions with means $\tilde{\mu}_i^r$ and μ_i^r , and standard deviations $\tilde{\sigma}_i^r$ and σ_i^r , such that $(\sigma_i^r)^2 = (\tilde{\sigma}_i^r)^2 + (\sigma_i^r)^2$, $\tilde{\mu}_i^r + (\tilde{\sigma}_i^r)^2 / 2 = 0$, and $\mu_i^r + (\sigma_i^r)^2 / 2 = 0$. We assume that private production forecasts, noted \hat{Q} , build on previous public forecasts if available:

$$(15) \quad \hat{Q}_{i,t}^r = Q_{i^r,t}^r \begin{cases} \exp\left(\sum_{j=i^r+1}^{i-1} \eta_{j,t}^r + \tilde{\eta}_{i,t}^r\right) & \text{if } i^r + 2 \leq i \leq i^r + 5, \\ \exp\left(\tilde{\eta}_{i,t}^r\right) & \text{if } i = i^r + 1. \end{cases}$$

Combining equations (5) and (15) gives $\exp\left(\tilde{\eta}_{i,t}^r\right) = \hat{Q}_{i,t}^r / \hat{Q}_{i,t}^r$: the size of the news shocks provided by USDA only can be obtained from the ratio of public and private projections. As before, we obtain that $\text{var}\left(\hat{Q}_{i,t}^r / \hat{Q}_{i,t}^r\right) \approx (\sigma_i^r)^2$.

This approach makes several simplifications. First, it assumes that private forecasts are immediately made public, which prevents assessing the benefits accruing to speculators from their privileged access to early information. Second, it assumes that there is one representative private forecast, while the reality is of a diversity of forecasts, which would give USDA forecasts a special role in decreasing the uncertainty related to

¹⁰This database was kindly provided to us by Scott Irwin.

the diversity of previsions. Third, and related, it considers that the only economic role of USDA forecasts is to improve the precision of information about fundamentals over a situation with only private forecasts. In the global games framework of [Morris and Shin \(2002\)](#), public information has another role: it allows agents with private information to coordinate while interacting strategically. Given the purpose of the article, these simplifications allow us to assess the benefits of public information compared to private information without the need for a model detailing the timing of each shock and without accounting for heterogeneous sets of information.

The theoretical framework assumes that public forecasts are more precise than private forecasts. So, a first step before decomposing the news shocks in their private and public elements is to verify that it is the case. [Table 3](#) displays the root mean square percentage error (RMSPE) for private and USDA forecasts using as final production value the January estimate. The forecasting error declines with time and is higher for private forecasts than for public forecasts.¹¹ This result confirms the model assumption. Using the ratio of public to private forecasts to identify the shock observed only by USDA, we cannot reject the null hypothesis that its mean is equal to 1. The standard deviation of this ratio gives $\hat{\sigma}_i^{US}$ from which $\tilde{\sigma}_i^{US}$ can be calculated by targeting the standard deviation of the total shock. Using this decomposition, the private forecasts predict between 61 and 77% of the total shocks variance.

Table 3. Informational Content of Private Forecasts about U.S. Harvest over 1970–2017

Month	RMSPE compared to January estimate			$\hat{Q}_{i,t}^{US} / \hat{Q}_{i,t}^{US}$		$\tilde{\sigma}_i^{US}$ (%) ^b
	Private (%)	USDA (%)	Difference (%)	Mean	SD (%) ^a	
July ^c	8.816	7.929	-0.887	1.000	1.897	3.123
August	6.312	6.207	-0.104	1.000	2.350	3.612
September	5.169	4.849	-0.320	1.004	1.738	3.181
October	3.280	2.677	-0.603	1.000	1.665	2.878
November	2.013	1.634	-0.379	1.002	1.081	1.351

Source: Authors' calculation based on forecasts provided by Scott Irwin.

Note: ^a Correspond to $\hat{\sigma}_i^{US}$. ^b Calculated as $\sqrt{(\sigma_i^{US})^2 - (\hat{\sigma}_i^{US})^2}$. ^c Based on data over 1983–2017 and calculated using USDA yield forecasts multiplied by private or USDA acreage forecasts.

Quantitative analysis

We now turn to analysis of the effects of news shocks.

Effects on market dynamics

We compare the model with news shocks described above, with a model without news shocks where the distributions of production shocks are degenerate and concentrated in the harvesting periods. [Table 4](#) presents

¹¹The results would be similar with a measure of mean absolute percentage error except for August when the private forecasts would be slightly more precise than the public forecasts.

the descriptive statistics on the asymptotic distribution of prices and stocks for the models with and without news shocks.¹² The changes in mean prices are not represented in the table because they are very small (below 0.3%). News shocks affect the distribution of prices without changing their means.

Table 4. Descriptive Statistics on the Asymptotic Distribution (All Results in Percentages)

Period	Shock	Prices					Stocks
		Standard deviation		Changes	Skewness	Kurtosis	Mean
		No news	News				
Oct./Nov.	σ_1^{US}	17.25	14.64	-15.1	-25.7	-41.7	-1.4
Nov./Dec.	σ_2^{LAC}	17.28	15.25	-11.7	-34.4	-49.8	-1.7
Dec./Jan.	σ_3^{LAC}	17.30	16.15	-6.7	-42.7	-58.5	-2.5
Jan./Feb.	σ_4^{LAC}	17.33	16.74	-3.4	-45.8	-62.3	-4.2
Feb./Mar.	σ_5^{LAC}	17.42	17.67	1.4	-47.5	-65.4	-14.8
Mar./Apr.	σ_6^{LAC}	15.43	14.35	-7.0	-19.0	-30.7	-0.9
Apr./May		15.46	14.38	-7.0	-19.0	-30.7	-1.1
May/June		15.48	14.40	-7.0	-19.0	-30.7	-1.3
June/July	σ_9^{US}	15.51	14.91	-3.8	-26.8	-38.6	-1.7
July/Aug.	σ_{10}^{US}	15.53	15.87	2.2	-34.7	-48.3	-2.4
Aug./Sept.	σ_{11}^{US}	15.56	16.83	8.1	-36.2	-51.6	-4.1
Sept./Oct.	σ_{12}^{US}	15.70	18.00	14.7	-31.7	-48.1	-15.7
All ^a		16.35	15.87	-2.9	-34.3	-51.8	-2.0

Source: Statistics calculated over 1,000,000 sample observations from the asymptotic distribution simulated with the model.

Note: ^a Statistics calculated over the pooled variables (prices or stocks) from all periods.

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 only a little with the seasons (1.99 percentage point maximum), increasing by about 0.03 percentage points between seasons except after new harvests when there are small jumps. These variations are easily explained. Between a harvest and the month immediately before the next harvest, there is no new information, and stocks are always positive, so from the non-arbitrage condition of storage (equation (1)) successive prices are deterministic functions of the harvest-time price: $P_{i+1,t} = (P_{i,t} + k\bar{P}) / \beta$.

The presence of news shocks changes seasonal price volatility a great deal. The differences between seasons increase with a maximum spread of 3.65 percentage points. In most seasons the volatility decreases significantly but not in February/March, July/August, August/September, or September/October; the volatility increases are the highest in the periods closest to the harvests (+14.7% in September/October preceding the U.S. harvest). As expected, between the seasons without news arrival, the same 0.03 percentage points increase in price volatility occurs as observed in the model without news shocks.

The relative changes in price volatility can be explained by the following mechanisms. Volatility decreases at harvest periods because the shocks received in these periods have a much smaller variance in the presence of

¹²See the online appendix for details of the numerical methods.

news shocks. Because of the intertemporal link created by storage, this reduction in price volatility leads also to price volatility reduction in the following periods. On the other hand, price volatility increases in growing periods before harvest, because that's when the news shocks happen. The volatility increase is the combination of two effects: the size of the news shock, the larger its variance the higher the price volatility increase, and its timing in the season. For the same variance, a news shock has a higher impact late in the season because stock levels decrease until the harvest and so the market is more sensitive to news late in the season when adjustment capacities are more limited. This explains why volatility increases most in September/October despite a news shock with a smaller variance than in the two previous months. These effects can be seen separately for the news about the U.S. and Argentina-Brazil harvests in table A2 in appendix. It shows that the news shocks about the U.S. harvest have more market impact than those about the Argentina-Brazil harvest, which is consistent with US news shocks that are more informative early in the season.

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 in the presence of speculative storage, prices are skewed positively (not represented in table 4): prices are concentrated below the mean with occasional positive deviations which possibly are larger than the negative ones. News shocks reduce the skewness considerably (between 19% and 48%) by reducing the occurrence of high prices significantly. Similarly, kurtosis of the price distribution is reduced by at least a third (although the distribution remains leptokurtic) by the introduction of news shocks because of the decreased probability of extreme price events.

It is striking that this reduction in seasonal price volatility and occurrence of high prices is obtained despite a reduction in mean stock levels. Without news shocks, it is profitable to keep higher levels of stocks before the harvest because there is a non-negligible likelihood of a bad harvest, thus, 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 stock levels in periods other than just before harvests seems less impressive because in these periods the mean stock levels are higher to smooth the bi-annual harvest over the whole year. However, in levels, not in proportions, the decrease is very similar for all periods. Stocks do not always reduce with news shocks. If early news shocks indicate a bad harvest, destocking is reduced compared to the situation without news which mitigates the price increase.

The last row of table 4 presents statistics on aggregate price volatility, when prices from all periods are pooled, and on mean stock changes for all periods. Price standard deviation decreases by 2.9%, showing that when the various periods are combined, the dominant effect of news shocks is to reduce price volatility. The very strong reduction in the higher-order moments of the distribution is related as before to smaller occurrence of price spikes. There is a small 2% decrease in mean stock levels. But, annual stocks usually are considered as the level at the beginning of the marketing year. Since in our model, the periods when there are non-zero probabilities of stockouts are February/March and September/October, the change in annual stock levels could be seen as the change in one of these seasons statistics: -14.8% or -15.7% .

To summarize the consequences of news shocks for market dynamics: one of the main effects of news shocks is to reorganize the price volatility across the seasons. They lead to a different allocation of seasonal price volatility with higher volatility just before harvests and lower volatility at other seasons. Overall price volatility is reduced, but its variation is small compared to the changes in seasonal volatility. News shocks

allow stocks to adjust before the harvest, reducing average stock levels, and the occurrence of price spikes.

Welfare effects of all news

Table 5 decomposes the welfare gains in the three efficiency components identified in equation (9): consumer efficiency gains, storage costs, and production costs. For all the news related to the two regions, having news shocks compared to having only one production shock at the harvest increases welfare modestly. Total gains amount to 0.38‰ of the value of annual steady-state consumption (column 1). To better understand the level of the welfare gains, we can compare them to the maximum gains attainable from having advance information about the coming harvests. For this we assume in column 2 that the total production is fully known one period after planting, and that there are no other shocks following this. In this case, total welfare gains equal 0.62‰ of the value of annual steady-state consumption. Thus, the actual news shocks are responsible for 61% of the potential gains attainable in the model.

Table 5. Welfare Changes (‰ of the Value of Annual Steady-state Consumption)

Welfare element	All information ^a				Public information ^b
	Actual news shocks (1)	Production known one period after planting (2)	US news shocks only (3)	LAC news shocks only (4)	US news shocks only (5)
Consumer's efficiency gains	0.07	0.12	0.03	0.02	0.01
Storage costs	0.35	0.55	0.18	0.16	0.05
Production costs	-0.04	-0.05	-0.03	-0.02	-0.00
Total gains	0.38	0.62	0.18	0.17	0.06

Note: ^a Comparison between the news shocks structure described in the column header and a situation without news shocks. ^b For the U.S. only, comparison between a situation with actual news shocks and a situation with only the shocks that would be observed by private analysts, $\hat{r}_{i,t}^{US}$; for Argentina-Brazil no news shocks.

The small size of the gains, potential and actual, especially compared to the important changes in price dynamics shown in table 4, is explained by the simplicity of the model which includes few channels through which price volatility could affect welfare. Consumers are not risk averse, and so are mostly indifferent to price volatility (Gouel, 2014). Similarly, in the model there are very few reasons why production costs should be affected by news shocks since production decisions are taken before observing news shocks, and mean prices change very little. So, average consumption is also little affected. The only significant source of welfare gains is related to storage costs. Advance information means that there is less need for high levels of stocks. The reduction in storage costs explains most of the welfare gains but given that storage costs are small compared to the value of consumption, used as a deflator, the eventual welfare gains are also low. That most of the welfare gains are related to the reduced inter-annual stock levels confirms the choice to use a storage model accounting for both inter- and intra-annual market dynamics rather than the approaches in Hayami and Peterson (1972) and Bradford and Kelejian (1977, 1978) which focus only on intra-annual storage and miss most of the effects of news shocks.

Given that most welfare gains are explained by reductions in storage costs, it is interesting to relate them

to the size of storage costs without news shocks. Under the same speculative motive, the model implicitly combines two types of stocks that we can label “deterministic” and “risky”. In our model where harvests occur twice a year but the commodity is consumed all the year, stocks have to be carried out for smoothing consumption between harvests (as in Samuelson, 1957). These stocks lead to “deterministic” storage costs as these costs would have been incurred in the absence of stochastic shocks in the model. These costs can be evaluated from the steady-state values of stock levels which are determined by assuming an absence of risk. News shocks won’t affect “deterministic” storage costs but only “risky” storage costs, the storage costs that are present in the model if there is some risk (calculated as average total storage costs minus “deterministic” storage costs). The total welfare gains from news amount to 2.31% of total storage costs and to 17.37% of “risky” storage costs. Thus, a sizable portion of the storage costs generated by uncertainty are reduced by the provision of news shocks. These expressions of welfare gains are consistent with the 2% reduction in average stocks and the 15.7% reduction in ending stocks observed in table 4.

It is also interesting to calculate the value of public information separately for each region, especially since we had to make more assumptions to identify the size of the news shocks for Argentina-Brazil than for the U.S. This is done in the columns 3 and 4 of table 5 by comparing a model where only one region has news shocks with the counterfactual without news shocks for any region. The welfare gains are very similar across regions, slightly smaller than half of the gains with both regions. The U.S. share of production of 42% should have commanded lower welfare gains than for Argentina-Brazil, but this is counterbalanced by the higher precision of the news shocks on the U.S. harvest.

To illustrate the welfare gains, we use a price of \$350 per ton of soybeans, and world production of 320 million tons based on recent data. World production then is valued at \$112 billion, and the total benefits from news shocks are valued at \$42.6 million.

Welfare effects of WASDE reports on U.S. harvest

The previous section analyzed the welfare effects of all news (private and public information and about the U.S. and Argentina-Brazil). We focus now on the welfare effects of public information coming in the WASDE reports, assuming that some information would be collected privately even in the absence of USDA and that it would be still publicly released as it is today. This can only be done for the shocks related to the U.S. harvest for which we have information about private forecasts. It is difficult in this case to choose the appropriate counterfactual, because private forecasts would be unlikely to stay the same in the absence of USDA forecasts. Private forecasters could increase their data collection and their accuracy, or on the contrary, reduce their accuracy because their forecasts partly build on public data that would be too costly to collect privately.

We have previously decomposed U.S. monthly shocks into a part observed by private analysts, $\tilde{\eta}_{i,t}^r$, and a part observed by USDA only, $\eta_{i,t}^r$. For the counterfactual with only private information, we assume that $\eta_{i,t}^r$ is not observed, which moves part of the shocks to harvest time when production is observable by all and we use the last column of table 3 to calibrate the standard deviation of private information, $\tilde{\sigma}_i^{\text{US}}$. Note that to maintain constant the standard deviation of the aggregate shock, the observability of production at harvest implies $\tilde{\sigma}_1^{\text{US}} = \sqrt{0.0769^2 - \sum_{i=9}^{12} (\tilde{\sigma}_i^{\text{US}})^2}$, which means that we cannot make use of the last row of table 3 related to the shocks at harvest.

Column 5 of table 5 presents the results in which a situation with the actual U.S. news shocks is compared to situations with only private shocks. For Argentina-Brazil, a situation without news shocks is assumed. These results allow us to estimate the contribution of USDA information accounting for the information provided by private analysts. The results can be compared with column 3 to estimate the share of the welfare gains related to private and public forecasts in total gains. The information uniquely provided by USDA accounts for one-third of the total welfare gains, consistent with the observation from table 3 that the private forecasts predict around two-thirds of the total shocks variance.

Using the same illustration as before, the welfare benefits of the USDA reports on the U.S. harvest, accounting for the presence of private forecasts, are \$6.7 million.

Improving the quality of advance information

In this section, we use the model to assess the effects of small improvements in the informational content of news shocks. These counterfactuals allow us to quantify what would be the benefits of improving the current crop forecasts (if it is possible). For example, by making available much more data about earth observation at higher frequencies, the CubeSat revolution is providing a basis for new approaches to crop production forecast, and in the recent years, new firms are relying on these big data to propose what they claim to be improved crop production forecasts. We consider four counterfactuals in which the last shock to the U.S. harvest has its standard deviation reduced by 1 percentage point to 0.73% by assuming that this part of the shock can be observed in one of the four earlier growing seasons. To maintain constant the standard deviation of the aggregate shock, $\epsilon_{1,t}^{US}$, the standard deviation of an earlier shock is increased according to $\sigma_i^{US} = \sqrt{0.0769^2 - \sum_{j \in \{1,9..12\}, j \neq i} (\sigma_j^{US})^2}$ for $i \in \{9..12\}$. The distribution means are adjusted to maintain unitary means for the production shocks.

Table 6 presents these counterfactual results. As can be expected, welfare gains are higher when the shock is moved to an earlier month. The maximum gains are achieved when moving the shock to η_9^{US} (column 3). Compared to the gains coming from news shocks on the US harvest (column 3 of table 5), the welfare gains amounts to an 4% increase. Regarding price volatility, table 6 presents the pattern described previously. Moving the information earlier in time moves price volatility when the information is received and slightly reduces the overall price volatility.

In these simulations, price volatility decreases in the first five months and increases after. The increases are limited (below 0.05 percentage points) before the period to which the shock is moved. The increase is more significant at the period to which the shock is moved and until the U.S. harvest. In column 3, price volatility increases by 0.15 percentage points in June/July when additional information is received. Price volatility increases similarly in the following months before the harvest. In column 6, the shock is just moved one month earlier. Price volatility increases by 0.02 percentage points before September/October and by 0.29 percentage points in September/October. These simulations show us that news shocks move the volatility earlier in times, when the information is received, and decrease it at harvest and later months because of a more efficient allocation of stocks. In consequence, news shocks about the U.S. harvest increase price volatility during the U.S. growing season and decrease volatility during the Argentina-Brazil growing

Table 6. Standard Deviation and Welfare Gains with Various Structures of News Shocks

Period	Shock	No news (1)	News (2)	Reduction of σ_1^{US} and increase of				
				σ_9^{US} (3)	σ_{10}^{US} (4)	σ_{11}^{US} (5)	σ_{12}^{US} (6)	
Standard deviation (%)								
Month								
Oct./Nov.	σ_1^{US}	17.25	14.64	14.34	14.37	14.41	14.49	
Nov./Dec.	σ_2^{LAC}	17.28	15.25	14.97	15.00	15.03	15.11	
Dec./Jan.	σ_3^{LAC}	17.30	16.15	15.89	15.91	15.94	16.01	
Jan./Feb.	σ_4^{LAC}	17.33	16.74	16.49	16.51	16.55	16.61	
Feb./Mar.	σ_5^{LAC}	17.42	17.67	17.44	17.46	17.49	17.55	
Mar./Apr.	σ_6^{LAC}	15.43	14.35	14.40	14.40	14.39	14.38	
Apr./May		15.46	14.38	14.43	14.42	14.41	14.40	
May/June		15.48	14.40	14.45	14.45	14.44	14.42	
June/July	σ_9^{US}	15.51	14.91	15.05	14.95	14.94	14.93	
July/Aug.	σ_{10}^{US}	15.53	15.87	16.02	16.04	15.90	15.89	
Aug./Sept.	σ_{11}^{US}	15.56	16.83	16.99	17.01	17.03	16.84	
Sept./Oct.	σ_{12}^{US}	15.70	18.00	18.19	18.21	18.24	18.29	
All ^a		16.35	15.87	15.83	15.83	15.84	15.85	
Total welfare gains (%)								
				0.383	0.398	0.396	0.393	0.389

Source: Statistics calculated over 1,000,000 sample observations from the asymptotic distribution simulated with the model.

Note: ^a Statistics calculated over the pooled prices from all periods.

season. Conversely, news shocks about the Argentina-Brazil harvest increase price volatility during the Argentina-Brazil growing season and decrease volatility during the U.S. growing season. When both sources of news shocks are combined, price volatility decreases everywhere except just before the harvests. These results help better explain that in table 4 the presence of news increases volatility more before the U.S. harvest than before the Argentina-Brazil harvest. The news about the Argentina-Brazil harvest reduces price volatility during the U.S. growing season, but since the news about the U.S. crops are quite precise the volatility tends to increase before the U.S. harvest and very significantly one month before. On the other hand, the precision of the news about the Argentina-Brazil harvest is much lower and the reduction of price volatility caused by the news about the U.S. harvest tends to dominate.

Sensitivity analysis

This section analyzes the sensitivity of the main results to the 3 key behavioral parameters: demand elasticity, supply elasticity, and storage cost. For each parameter, we consider two scenarios: a halving and a doubling of the benchmark parameter, which should provide reasonable lower and upper bounds of the likely values. Table 7 presents the results of the sensitivity analysis with for parsimony a focus on the most important results.

To help the comparison, column (1) reproduces the results under the default parameters.

Table 7. Sensitivity Analysis

	Default	Demand elasticity		Supply elasticity		Storage cost	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
α^D : Demand elasticity	-0.2	-0.1	-0.4	-0.2	-0.2	-0.2	-0.2
α^Q : Supply elasticity	0.4	0.4	0.4	0.2	0.8	0.4	0.4
k : Storage cost (% of \bar{P})	0.56	0.56	0.56	0.56	0.56	0.28	1.12
Price SD in Sept./Oct. under news (%)	18.00	28.16	11.29	19.31	16.75	14.89	22.10
Change in price SD in Oct./Nov. (%)	-15.1	-22.0	-16.4	-16.2	-13.6	-14.6	-17.3
Change in price SD in Feb./Mar. (%)	1.4	-5.6	-2.7	-0.8	4.7	2.4	-2.2
Change in price SD in Sept./Oct. (%)	14.7	13.7	7.9	13.2	16.9	16.0	10.1
Change in all prices SD (%)	-2.95	-8.00	-6.40	-4.33	-1.04	-2.20	-6.19
Ratio of total welfare gains to storage costs (%)	2.31	1.96	2.30	2.66	1.91	2.28	2.27
Ratio of total welfare gains to risky storage costs (%)	17.37	10.97	29.59	19.85	14.42	13.53	24.23

In the first row after the parameters, we can first verify with the price standard deviation in September/October (the month with the highest volatility) under news that the parameters have the expected effect. The results are as expected. Demand elasticity has a strong influence on price volatility with an increase to 28% if $\alpha^D = -0.1$ and a decrease to 11% if $\alpha^D = -0.4$. Supply elasticity has a smaller effect since supply takes time to adjust. Lastly, a higher storage cost makes speculation more costly, so storers buy at lower prices and sell at higher prices leading to more price volatility.

Three changes in price volatility after the introduction of news are reported in table 7: October/November, February/March, and September/October, corresponding respectively to the U.S. harvest when the decrease in the volatility is the highest, the period before the Argentina-Brazil harvest when the volatility increases slightly in the benchmark, and the period before the U.S. harvest when the volatility increases the most. For the two extremes, the same pattern holds with a strong decrease in October/November and a strong increase in September/October. However, in February/March, the sign of the change depends on the parameters' value. Remember that there are two effects at play: the news shocks on the U.S. harvest decrease volatility during the Argentina-Brazil growing season and the news shocks on the Argentina-Brazil harvest increase it. Depending on the respective strength of the two effects, volatility may increase or decrease in February/March.

The results on overall price volatility are consistent across the parameterizations: news shocks tend to decrease price volatility.

Lastly, regarding the welfare gains, they are quite stable around 2% if expressed in ratio to storage costs. In ratio to risky storage costs, there are more variations because in this case both the numerator and the denominator may vary significantly. Since this expression of welfare tends to be related to the decrease in the level of the end-of-season stock, it shows that the presence of news shocks has a robust effect on decreasing end-of-season stocks.

Conclusion

In this article which uses the example of the global soybean market, we have shown how news shocks that provide advance information about future production of a storable commodity affect market dynamics,

compared to a situation where production is known only when it is realized. The results were obtained by building a rational expectations storage model able to account for intra-annual and inter-annual dynamics, and running counter-factual simulations. The monthly crop production forecasts, from USDA and private firms, were used to calibrate realistic news shocks in the model. These shocks matter greatly for market dynamics because based on the information they provide, storers adjust their stocks before the harvest to be more consistent with the newly expected market tightness. For example, if a poor harvest is forecast, storers reduce their destocking before the harvest which in turn reduces the occurrence of price spikes. The presence of news shocks changes price volatility a lot, decreasing it at harvest and after and increasing it during the growing season when news arrive. Volatility increases the most in the periods just before harvest, because the low levels of stocks in these periods make the market very sensitive to new information. Price volatility slightly decreases overall, but the main effect of news shocks is to redistribute price volatility during the season.

The possibility to adjust stocks early in the season based on advance information reduces the need to carry large inter-seasonal stocks. The reduction in stock levels allowed by news shocks explains most of the welfare gains from having advance information which amount to 2% of the value of storage costs, when considering all news shocks from USDA and private firms. This corresponds to welfare benefits of news of \$43 million for news about the U.S. and Latin America harvest, \$20 million being related to the U.S. harvest. Since for the U.S. harvest, seasonal crop forecasts by private firms predict two-thirds of the total shocks variance, they also account in the model for most of the welfare benefits of news shocks, the increased precision brought by USDA reports accounting only for \$7 of the \$20 million.

This study has dealt with only a small part of the informational issues that occur in commodity markets, and thus, with only a small part of the potential benefits of public information in these markets. For example, there are problems related to reporting stocks. Stock levels are not directly observable by market participants but are reported regularly by government agencies and international organizations. These statistics although crucial for the market are recognized as not wholly reliable: the preciseness of their estimation is questionable since they are based either on surveys, and more frequently outside the U.S., as residuals in commodity balances. In our storage model framework, this means that agents do not observe the state variables with precision, and have to make their decisions based on incomplete information. The reporting of stock levels is another, and perhaps as economically important problem as advance information on the coming harvests. Another feature of our approach that might have led to underestimation of the benefits from public information is that we developed a simple model with few margins for adjustment to new information. An extension to our model that includes agents taking costly decisions during the growing season would increase the benefits: for example, planting crops that are close substitutes or buying livestock feed.

Beyond the question of the cost/benefit analysis of public information, this research may have other policy implications related to the opportunity for price stabilization policies. Following the 2007/08 food crisis, we observed renewed interest in food price stabilization policies (Gouel, 2014), in particular to protect consumers from price spikes. These policies frequently involve the combination of storage and trade policies which can be very costly fiscally (Gouel and Jean, 2015). The results in this article show that provision of public information during the growing season is effective to reduce the occurrence of price spikes without the need for public interventions in the market. Since market information systems in developing countries are

generally less developed than in the U.S., their development should be considered before considering any price stabilization policies.

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. (2012). Quantifying the WASDE announcement effect. *American Journal of Agricultural Economics*, 94(1), 238–256.
- Adjemian, M. K., Johansson, R., McKenzie, A. and Thomsen, M. (2018). Was the missing 2013 WASDE missed? *Applied Economic Perspectives and Policy*, 40(4), 653–671.
- Beaudry, P. and Portier, F. (2014). News driven business cycles: Insight and challenges. *Journal of Economic Literature*, 52(4), 993–1074.
- Bobenrieth, J. R. A., Bobenrieth, E. S. A., Wright, B. D. and Zeng, D. (2017). “Negligible” Trends, Spurious Linearity. mimeo.
- 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.
- C-FARE (2016). *From Farm Income to Food Consumption: Valuing USDA Data Products*. Report, Council on Food, Agricultural & Resource Economics.
- Chambers, M. J. and Bailey, R. E. (1996). A theory of commodity price fluctuations. *The Journal of Political Economy*, 104(5), 924–957.
- Egelkraut, T. M., Garcia, P., Irwin, S. H. and Good, D. L. (2003). An evaluation of crop forecast accuracy for corn and soybeans: USDA and private information agencies. *Journal of Agricultural and Applied Economics*, 35(1), 79–95.
- 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.
- Good, D. L. and Irwin, S. H. (2006). *Understanding USDA Corn and Soybean Production Forecasts: Methods, Performance and Market Impacts over 1970–2005*. Project Research Report 2006-01, AgMAS.
- 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. and Jean, S. (2015). Optimal food price stabilization in a small open developing country. *World*

- Bank Economic Review*, 29(1), 72–101.
- 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.
- IMF (2014). *World Economic Outlook: Recovery Strengthens, Remains Uneven*. Washington, DC: International Monetary Fund.
- Isengildina, O., Irwin, S. H. and Good, D. L. (2006). Are revisions to USDA crop production forecasts smoothed? *American Journal of Agricultural Economics*, 88(4), 1091–1104.
- 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.
- Lowry, M., Glauber, J. W., Miranda, M. J. and Helmsberger, P. G. (1987). Pricing and storage of field crops: A quarterly model applied to soybeans. *American Journal of Agricultural Economics*, 69(4), 740–749.
- Merener, N. (2015). Globally distributed production and the pricing of CME commodity futures. *Journal of Futures Markets*, 35(1), 1–30.
- Morris, S. and Shin, H. S. (2002). Social value of public information. *The American Economic Review*, 92(5), 1521–1534.
- Ng, S. and Ruge-Murcia, F. J. (2000). Explaining the persistence of commodity prices. *Computational Economics*, 16(1–2), 149–171.
- Nordhaus, W. D. (1987). Forecasting efficiency: Concepts and applications. *The Review of Economics and Statistics*, 69(4), 667.
- Osborne, T. (2004). Market news in commodity price theory: Application to the Ethiopian grain market. *The Review of Economic Studies*, 71(1), 133–164.
- 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.
- Samuelson, P. A. (1957). Intertemporal price equilibrium: A prologue to the theory of speculation. *Weltwirtschaftliches Archiv*, 79, 181–219.
- Steinwender, C. (2018). Real effects of information frictions: When the States and the Kingdom became united. *The American Economic Review*, 108(3), 657–696.
- 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.