UNIVERSITY OF CALIFORNIA

Los Angeles

Fundamentals, Speculation, and Seasonal Correlation in Commodity Markets

A dissertation submitted in partial satisfaction

of the requirements for the degree Doctor

of Philosophy in Economics

by

Francisco Arroyo Marioli

2019

ABSTRACT OF THE DISSERTATION

Fundamentals, Speculation, and Seasonal Correlation in Commodity Markets

by

Francisco Arroyo Marioli

Doctor of Philosophy in Anthropology

University of California, Los Angeles, 2019

Professor Pierre-Olivier Weill, Chair

The understanding of agricultural commodity financial markets has become of significant interest, given the increasing attention both the private and public sector have been giving to them. Financial investment in these goods has increased exponentially over the past 15 years. Food prices have reached significantly high levels. Therefore, my dissertation focuses on what is a relevant not only for the literature but also for both public and private institutions. I start by working within a competitive storage framework, as is usual for the literature. I then make assumptions and change the timing and information structure to match realistic aspects, therefore obtaining partially different theoretical results. I then intend to test these results by contrasting them with publicly available data regarding prices, production, consumption and information. Since the model is designed to match real-life aspects of the market, it can be applied to identify and measure fraction of price changes due to each different fundamental. For example, one recurring and important aspect of the data is that, in general, storage models predict excessively stable prices. That is, standard deviations are higher in the data than those compared to simulations in the models. Volatility is important since it can have potential welfare effects on both consumers and producers. Therefore, it is an issue that deserves attention. Moreover, related to this, in the past 15 years financial

investment in these markets has increased severely, bringing concern to policy makers since these may have some effects on price levels and/or volatility. In the first chapter, I propose an innovative structure in the model to study this. More specifically, I subdivide time periods in four quarters, and each quarter with its own specific parameters. That is, only in the first quarter there is production, and demand presents seasonal effects for each of the four quarters. My intention is to improve the accuracy of the model by introducing once more a more realistic framework. Once these adjustments are made, I will be able to decompose and quantify through simulations the different causes of prices changes.

In the second part, I incorporate an innovation into the standard theoretical sotrage model. The cornerstone of seasonally produced goods literature is the competitive storage model. Since production occurs only during one part of the year but consumption takes place all year along, inevitably storage appears as the main solution. Therefore, storage models have been widely used within the literature, with an important deal of success. However, not all aspects of the data have yet been explained. For instance, when it comes to agricultural goods, the model predicts that future contracts that deliver goods before the next harvest should not be strongly correlated with futures that deliver goods after the harvest takes place. The argument for this is that the first contracts deliver goods "from last year", whereas the latter ones deliver "this year's harvest". Since sources of supply are different, when new news regarding supply appear (for example, a harvest forecast) they should only affect the latter contracts, but not the first ones. The data shows however otherwise. Indeed, correlation between "new harvest" and "old harvest" futures contracts is positive and close to 1. This is the issue I address in the second chapter. The key element in my paper is that I assume that harvest comes in "continuously" within a relevant time interval instead of "all in one moment". This allows me to split the harvest between early and non-early parts. I show that the market equilibrium results in the early part end up being arbitraged with "old" future contracts, whereas the non-early section arbitrages with "new" ones. Therefore, the same source of supply gets sold on both type of contracts, allowing for supply induced positive correlation. I simulate the model and show this result is robust to changing parameter specifications, obtaining correlations between 0.7 to 1, as in the data. I also provide proof of the assumptions made to get this result, showing that they are highly realistic. These results are not incompatible with the main findings that have already been made, therefore it contributes to the literature by additionally explaining an unsolved puzzle.

In my third chapter, I analyze inflationary processes in major LATAM economies. More specifically, with other two coauthors we study inflation in Peru, Colombia, Brazil, Mexico and Chile for the past 18 years. We find that domestic factors such as intertia and expectations still play the biggest role. Foreign inflation however gains importance in some countries. With regarding to Phillips curve slopes, we find that these have been flattening in the last decade for most countries, that is, the cycl has a smaller effect than it used to have in previous decades when determining inflation.

The dissertation of Francisco Arroyo Marioli is approved.

Mathew Saki Bigio Luks

Gary Richardson

Ariel Tomas Burstein

Stavros Panageas

Pierre-Olivier Weill, Committee Chair

University of California, Los Angeles

2019

Contents

List of figures and Tables

Trading Places: The role of Agricultural Commodity Fundamentals, Information and Speculation

Commodity Futures and Seasonal Correlation

Drivers of Inflation in LATAM

Acknowledgments

Chapter 3 is a version of the paper that is currently in preparation for publication with Vibha Nanda and Frederik Toscani.

VITA

Francisco Arroyo Marioli

EDUCATION

PhD. (c) in Economics, UCLA	2014–2019 (Expected)
M.A. in Economics, UCLA	2015
M.A. in Economics, Torcuato DiTella University	2013
B.A. in Economics, University of Buenos Aires	2011

RESEARCH IN PROGRESS

Trading Places: The role of Agricultural Commodity Fundamentals, Information and Speculation. Job Market Paper.

Commodity Futures and Seasonal Correlation.

Liquidity, Networks, and Unintended Consequences: The Founding of the Fed and the Great Depression. Coauthored with Carrera F. and Richardson G.

Revisiting Phillips Curves in Latin America. Coauthored with Nanda V. and Toscani F.

HONORS & FELLOWSHIPS

Torcuato DiTella University Economics Department. Graduate Tuition Fees Waiver Fellowship.

UCLA Economics Department. PhD Fellowship.

University of Buenos Aires. Magna Cum Laude – Bachelor's Degree.

RESEARCH EXPERIENCE

Research Assistant, Torcuato Di Tella University (2013-2014).

PRESENTATIONS

"Exchange Rate pass-through, the Brazilian and Mexican case", coauthored with Pascuini P., Argentine Association of Political Economics, Trelew, Argentina, 2012.

TEACHING EXPERIENCE

Instructor

Torcuato Di Tella University. Undergraduate. Macroeconomics (January 2014).

UCLA. Undergraduate. Money and Banking (Fall and Spring, 2018).

Teaching Assistant

University of Buenos Aires. Undergraduate. Macroeconomics (2009 – 2012).

Torcuato Di Tella University. Graduate and Undergraduate (2012 -2014).

UCLA. Undergraduate. Econ 1 & 2. Advanced Macroeconomics (2015 - 2018).

Part I

Trading Places: The role of Agricultural Commodity Fundamentals, Information and Speculation

Francisco Arroyo Marioli, UCLA

What explains the surge and plunge commodity markets have undergone in the past 20 years? Are speculators to be blamed? Do prices reflect full information? These are the main questions I address in this paper, in the context of the corn market. I formulate and calibrate two quantitative models of corn prices formation. The first model is designed to explain prices in the long run (annual frequency), while the second model applies to prices in the short run (quarterly frequency). For the long-run analysis, I find that deviations of theoretical prices from observed ones are very small after 1996, and before 1996 they can be explained by government intervention. For the short-run analysis, my model is designed to mimic the typical seasonality seen in agriculture markets, incorporate supply and demand shocks as well as news shocks, and allows for speculative storage decisions. I find that demand and supply fundamentals can account for around 52% of past price changes from 1975 to 2016. I also estimate the impact of information shocks to explain an additional 18% of quarterly deviations. Finally, find that at least 30% of short-run price changes seem to have explanations other than supply or demand fundamentals or information, demonstrating that when analyzing quarterly data, prices do not always closely track fundamentals.

1 Introduction

The objective of this paper is to study the fundamental and non fundamental determinants of commodity prices over the past decades. The analysis specifically focuses on corn markets, for several reasons. Corn and soybean are the main agricultural commodities in terms of production market value. They are the most traded contracts in futures market (next to cotton and wheat). Finally, the US is world leader in corn production, consumption and exports.

In this paper I compare prices defined by fundamentals versus observed prices. I define fundamental determinants of price as demand and supply shocks. These may be either current or future. Future demand and supply shocks affect current prices through information shocks (agents acknowledge that shocks will occur in the future) that change current inventory holding decisions. I divide this paper into two parts. In the first part, I study the current determinants of corn prices at an annual frequency, while in the second I study the short-run, cyclical determinants. The first section determines how annual prices would have changed solely based on current demand and supply shocks. The reason for this is simplicity: I show that even when focusing only on current shocks, the model fits the data very well. Any additional feature would only increase the performance of the model, thus making my results only stronger. In the second part, I construct theoretical prices at a quarterly frequency. Given that supply is zero in some quarters, I introduce inventory purchases as part of the market. Given that inventory purchases are made within a profit maximizing scheme, expectations, and therefore future shocks through information, start playing a key role. I hence introduce in this section future shocks as part of fundamental-determined prices.

More specifically, in the annual model, I analyze the long-run price trend from 1975 to 2015. I then simulate a theoretical price time series, defined as the price such that each year's demand matches supply or, in other words, a price such that inventory variations are zero. Moreover, it is even possible to assume that inventories could reach zero at the end of the year, given that, in practice, in the last quarter it usually reaches very low values with respect to harvest size. I proceed to compare these theoretical prices with observed ones and estimate the difference, and find that after 1996, differences are small and do not last more than a year or two. Additionally, I find that the increase in price levels in the 2000s seems to be explained mostly by the presence of an increasing demand for bioethanol.

In the quarterly model, I design a model with a market structure that mimics the typical seasonality seen in agriculture markets (with production occurring only during one quarter). Given that the time lapse between decisions is shorter, I also allow for information shocks to occur by introducing them to the agent's expectations through private and public signals. Results show that around 70% of price volatility can be explained through the shocks cited above, leaving a 30% measured space for other non fundamental sources of variation.

Regarding model specifics, in the long run case I assume isoelastic forms $D_t = Z_t p_t^{-}$ for demand and $S_t = A_t p_t^{\eta}$ for supply, the first term represents (exogenous) levels and the second term the endogenous response to price with the respective elasticity. Therefore, shocks are captured as changes in Z_t, A_t. These will be defined as the *current fundamental* source of price changes. Time series for these variables can be obtained using USDA corn usage and production data. They are then integrated into a market clearing equation with no storage decisions. In the short run case I consider demand as a whole aggregate and assume identical isoelastic functions for both demand and supply, as before. I linearize the model, calibrate it and simulate with real shocks as inputs. I then estimate the impact of different shocks (demand, supply, and information) on price changes and quantify them.

One additional analysis that results from this paper is related to the issue of prices' being fully informative and appropriate for business-cycle measurement. Romer (2006) raises this question and shows evidence of firms that not always behave in profit-maximing ways by analyzing the case for professional football teams. In this paper I show that although in the long run, prices seem to be very close to fundamentals, when going to a quarterly frequency, market prices can sometimes be very far away—as much as 50%—from them. These results are in line with Hussman (1992), who shows that once imprecise signals are introduced in a rational expectations framework, market prices become inefficient in transmitting information. Because corn markets are considered to be well-functioning markets, in the sense that prices and transactions are transparent, centralized, and very liquid, it is surprising and interesting to find that they might not always behave as one would expect in a classical supply-demand model with inventories and utility and profit-maximizing agents. This raises several relevant questions regarding macroeconomics, financial markets and industrial organization. First, if a very well developed market fails to deliver fully informative prices, what can then be said regarding other markets, where illiquidity, information asymmetry and search costs are more relevant? Moreover, since these results are found for quarterly data, the implications for business-cycle accounting could be important. In several developing countries commodity markets such as corn have a significant impact in overall GDP and exports. If the prices that are being used for measurement are not market-clearing ones, important distortions could be taking place. Also, since annual prices don't differ from fundamentals as much as quarterly ones do, another potential question arises: Should we produce national accounts using prices measured at quarterly, annual, or some other frequency? How much time do markets need to become fully informative? Many business-cycle theorists believe that shocks on the real side of the economy, such as shocks to TFP and commodity prices, trigger and propagate economic fluctuations. Commodity price shocks clearly play important roles in developing nations today, particularly resource exporters and have also played a large role in economic fluctuations in the past. Economists and policy makers believed, for example, that the commodity-price declines in the late 1920s and early 1930s contributed to the length and severity of the Great Depression. In 1933, the Roosevelt administration's efforts to raise commodity prices, particularly prices of farm products like corn, formed the centerpiece of its efforts to resurrect the economy. My results have implications for the sources of shocks and the accuracy of the interpretation of the *shocks* derived from RBC models. Given that productivity is a key driver of the cycle, a price system that does not track fundamentals closely could result in inefficient resource allocations, or in other words, aggregate productivity losses. Therefore the implications of these findings are relevant regarding short-term business cycles. In the short run (at business-cycle frequencies), the majority of commodity price shocks do not reflect supply and demand, but could instead be affected by speculative factors. These distortions could have a significant impact on commodity producing economies.

This paper is divided into four sections. The first section introduces and explains the main structure of the paper and its relation with the literature. Section 2 presents the low-frequency model and its results. Section 3 presents the high-frequency model, calibration, and estimation results. Section 4 summarizes and concludes.

1.1 Literature Review

This paper, therefore, contributes to the literature in three respects: first, it tests the hypothesis that the spike in commodity prices was due to by non fundamental reasons, but focuses on a less explored market, since most of the literature has focused on oil markets. Second, I add a new feature to high-frequency analysis: information shocks. In this paper, I am able to quantify the historical impact of information shocks by incorporating USDA reports as a source of information. Additionally, I estimate private signals and their impact on prices, showing that when it comes to quarterly analysis, they are a relevant source of volatility. Third, I decompose price changes per source of change—that is, I estimate the impact of each variable change each year, from 1975 to 2015, and therefore offer alternative explanations for the observed change in prices in past years. This also allows me to estimate the fraction of prices that cannot be explained by fundamental or information factors, leaving a measured space for further research.

The literature has also examined whether commodity markets have been altered for non fundamental factors in the past decades, such as financial speculation. For instance, in energy markets results tend to indicate small or null effects of speculation on oil prices. Kilian and Murphy (2013) develop a VAR model with speculative demand shocks and contrast it with recent oil inventory data and find no basis for speculation's being blamed for the 2003-2008 price period. They do, however, find it plausible that there was some influence in previous years 1979, 1986, and 1990. Knitell and Pyndick (2016) analyze the oil market using a simple static partial equilibrium model with inventory markets and again find no relation between speculation and the oil price peak in 2008. Fattouh, Kilian, and Mahadeva (2013) summarize the literature that examines oil markets and conclude that there is no evidence that speculation is the main driver of price increases.

Another branch analyzes commodity financial markets from a portfolio point of view. Bohl and Stephan (2012) study the effect of increased trading in future markets for six top traded agriculture and energy commodities using a GARCH model approach and find no evidence of a causal relation between future trading and price volatility. Chary, Lochstoer, and Ramadorai (2013) show that restrictions in financial markets can alter spot prices through hedging decisions, affecting real outcomes. Sockin and Xiong (2015) study on the possibility of information frictions in commodity markets, and demonstrate the importance of prices as signals for both demand and supply and the weakness of assuming that shocks are publicly known. In line with their results, this paper incorporates information shocks and allows for some shocks to be unknown.

The work presented in this paper is mostly related to Knitell and Pyndick (2016), who estimate non fundamental shocks as deviations in inventory levels. I follow a similar methodology except that I consider corn markets (instead of oil) and a dynamic framework (rather than static).

2 Long Run Model

The goal in this section is to analyze and explain changes in price levels. In particular, one key question is whether observed prices have deviated from fundamentals because of the presence of non fundamental factors, such as, due to the growing importance of financial speculation. I propose the following experiment: Calculate the theoretical price that would balance supply and demand every year and compare it with the observed price. That is, calculate the price such that inventory variation would have been zero. The intention behind this simple experiment is to see how prices would have changed solely based on *current* demand and supply shocks.Any additional features added here would only enhance results, therefore the fit that I will show can only be enhanced but not worsened by any additional feature of analysis we might want to add. I will show that even under this simple framework, results are very conclusive. It also implies an more long-run based approach. In the long-run, inventories cannot play a significant role in price determination. Under the plausible assumption that speculators do not seek to buy-and-hold inventory, it is clear that in the long run prices cannot deviate far from fundamental ones. To understand this better, assume it is not the case. That is, assume that the fundamental price is systematically below or above equilibria with no inventory change. This would imply that stocks either decrease every year or accumulate infinitely. Neither scenario is consistent with long-run equilibria, since it would either imply hitting the zero lower bound or an irrational accumulation of stocks. Since the analysis is long-run based, I believe this is a useful approach.

Formally, the theoretical price results from the following market-clearing equation:

$$Z_t p_t^{-\omega} = A_t p_t^{\eta}. \tag{1}$$

$$Z_t = Z_t^{food} + Z_t^{feed} + Z_t^{ethanol} + Z_t^{\text{exports}}$$

The left-hand side of the equation is demand for corn, and has four subcomponents: ethanol, food industry, feed (and residual) and exports. Supply has only one source: farming. Imports in the US are (practically) zero. For simplicity, I assume that demand and supply are isoelastic functions of the price, as in Knitell and Pyndick (2016) and similar to Deaton and Laroque (1996). In this model, for demand functions, Z_t, captures exogenous demand shocks and $p_t^{-\omega}$ the endogenous response to prices, with ω being the demand elasticity parameter. For supply, A_t represents productivity and η supply elasticity.

Equation (1) thus defines the theoretical market price as $p_t^* = \left[\frac{Z_t}{A_t}\right]^{\frac{1}{\omega+\eta}}$. Also, to quantify the impact of different shocks, I can calculate the first-order effect for each of them.

$$dp_t = \sum_i \frac{\partial p_t}{\partial Z_t} dZ_t^i + \frac{\partial p_t}{\partial A_t} dA_t, \quad i \in \{\text{food,feed,ethanol,exports}\} \tag{2}$$

where the right-hand side gives us the sum of the effect of changes in all $i \in \{\text{food,feed,ethanol,exports}\}$ demand fundamentals and production fundamental A.

USDA databases have time series for prices, all four demand uses (ethanol, feed, food, and exports), production, and yield per acre. I use these to estimate elasticities and identify exogenous shocks. For demand price elasticities, I use instrumental variables, taking yield per acre shocks as a first-stage instrument for prices. Changes in yield per acre are explain mostly by weather factors. Farmers have almost no control over short run productivity levels. Therefore, it is reasonable to assume that they are exogenous to the production process and are appropriate instruments for prices. To estimate supply elasticity, I take estimations from Roberts & Shlenker (2013). Estimation results can been seen in Table I.A.. Results fall within other literature's findings[1]. Once elasticities are estimated, given that I have demand, supply and price data, I can identify exogenous components by reversing the isoelastic equation and setting the exogenous component equal to the demand or supply level times prices elevated to inverse elasticity.

[1] The Food and Agriculture Policy Institute shows elasticity estimations for corn in different countries, giving results that are always between 0.1 and 0.5. for both demand and supply. http://www.fapri.iastate.edu/tools/elasticity.aspx

$$
\begin{aligned}
Z_t^i &= D_t^i p_t^{\omega}, \\
A_t &= S_t p_t^{-\eta}.
\end{aligned} \tag{3}
$$

Given time series for A_t, Z_t^i, and p_t, and estimated values for ω, η, I can decompose price changes per year by estimating the different terms of equation (2). As a result, changes in inventories explain the remaining residual. Results per variable can be seen in Figures I to V.

I solve the price-solving equation (1) from 1981 to 2015. Figure I shows the theoretical price that results from the simulation and compares it against observed prices. The figure suggests two key observations: First, differences between the model-based and observed price are much larger before 1996 than after. That year, a more market-friendly US agriculture policy bill was passed, which reduced the budget for government purchases (hence the scope for intervention) and lowered price floor targets below market equilibria. It is clear that after such an event, the distance between both time series reduces heavily. Regarding the 2000s spike in prices, the hypothesis of a non-fundamental driver seems pretty weak, since prices moved as one would expect them given demand and supply shocks. Second, theoretical volatility reduces significantly after the year 1996.

Table I.A

Description	Parameter	Value
Demand Elasticity	ω	0.18***
Supply Elasticity	η	0.15***

Table I. *,**, and *** indicate p-values inferior to 0.1, 0.05, and 0.01, respectively.

Table I.B

Variable	Standard Deviation		
	All	Prior to 1996	After 1996
Observed Price (in logs)	0.39	0.36	0.32
Fundamental Price (in logs)	0.59	0.74	0.36
Observed to Fundamental Ratio	0.41	0.59	0.12
	Correlation		
Observed vs Fundamental Prices	0.82	0.72	0.96

Table I.B.The first two rows show standard deviations of both observed and fundamental prices for each time interval. The row "Observed to Fundamental Ratio" indicates the standard deviation of the ratio of observed prices divided by fundamental ones. The final row "Observed vs Fundamental Prices" shows the correlation between observed and fundamental prices for each time interval. A traditional statistical F test was performed to check for a null hypothesis of equality in standard deviations of the ratio before and after 1996, results rejected the null hypothesis at a 1% level of significancy.

Figure I

Theoretical Prices vs Observed

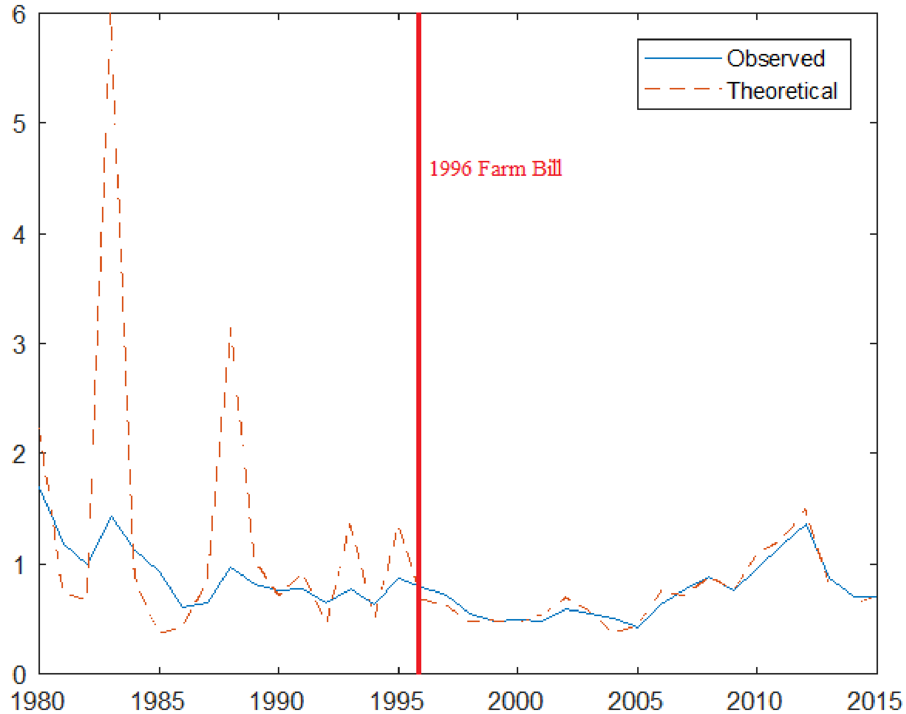


Figure I. Theoretical prices are the values that solve the market clearing equation (1) given the identified shocks. Observed ones are those obtained from the data. Units are in dollars per 100 metric tons of corn.

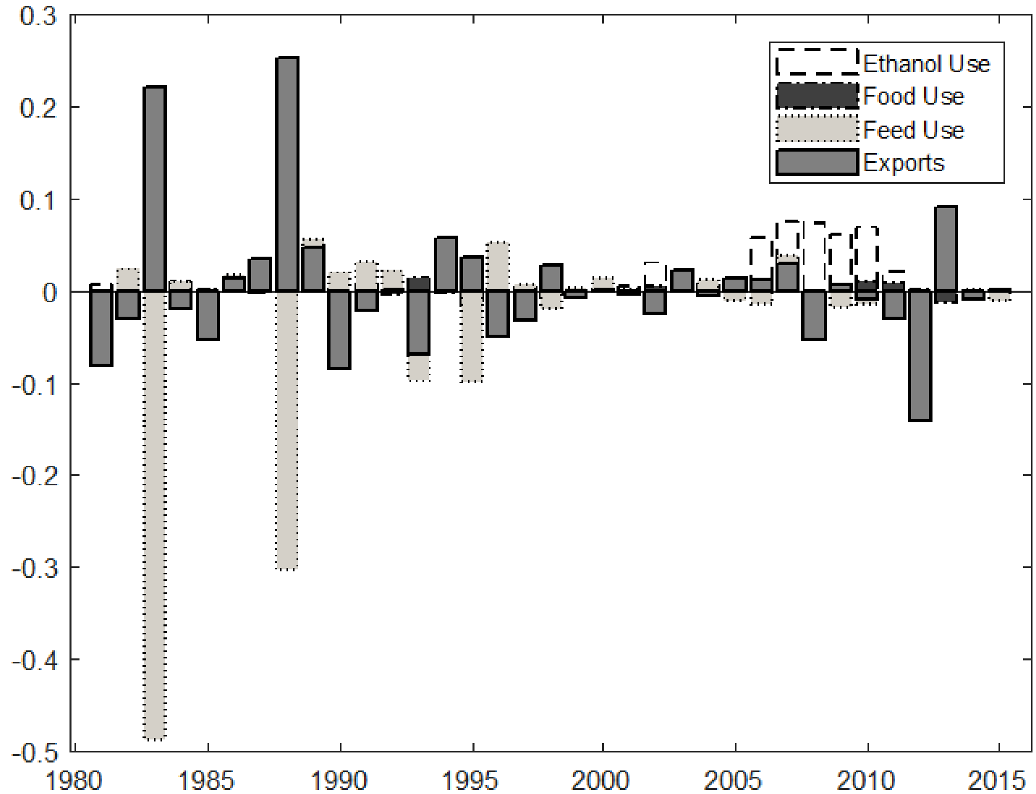
In 1996, a Farm bill was passed. That significantly shifted U.S. farming policy from a highly interventionist one toward a more "free market" approach. Prior to this bill, U.S. farming policy was heavily biased toward sustaining minimum price levels set at the discretion of the policymaker. The main tool through which this took place was by either government purchase of goods (The Commodity Credit Corporation - CCC- program) or subsidies to storage in the private sector (Farmer-Owned Grain Reserve program). Figure IV quantifies the effect these programs had on price levels. After 1996, the government budget allocated for purchasing goods was minimized, and minimum prices were set below market values; this resulting in a practically zero direct government intervention in corn markets. Table I.B shows statistical moments before and after the bill was passed in 1996. Table I.B shows standard deviations for observed prices and fundamental ones, standard deviation of the ratio and correlation. Two important facts emerge. First, the ratio between observed price and "fundamental" ones becomes significantly more stable after 1996. A traditional statistical F test was performed to check for a null hypothesis of equality in standard deviations of the ratio before and after 1996, with results rejecting a null hypothesis at less than 1% significance. This is also in line with an increase in correlation between both series. Second, the volatility of fundamental prices themselves also drop after 1996. This new stability could perhaps be one of the reasons for the lack of direct intervention during those years.

Regarding the 2000s, it is clear that observed prices do not deviate too significantly from fundamental ones. That is, the price that would have theoretically cleared the market with no storage decisions has not deviated significantly from observed prices. This suggests that speculation (or other non fundamental shocks) has had little—if any—effect on market prices.

The immediate question is then: If not speculation, what has drives the price spikes? Apparently, as can be seen in figure II, ethanol explains almost 56% of the price increase between 2005 and 2010. This can partially be explained by technological improvements (that allow ethanol to be used in energy markets) and government policy that has forced gasoline producers to incorporate mandatory fractions of ethanol in their final products. This decomposition has been calculated by inputting the estimated values of fundamentals and elasticities into equation (2). Additional evidence for this is the increased correlation between oil and corn prices, as in Figure VII (see Appendix). Given that an increasing fraction of corn is used for ethanol, increasing correlation with energy markets is therefore to be expected, and this is what effectively is seen in the data. This can be viewed as a new form of intervention, since instead of directly purchasing the product, they force private agents to use them. In future work, it would be interesting to use these results and calculate the impact of this intervention compared to previous ones, as well as other policy implications.

In conclusion, the model used here suggests that there is very little ground for a nonfundamental explanation of price changes in levels during the past 20 years. More specifically, regarding the 2000s, the impact of biofuels on corn markets seems to be one of the main explanations for the observed behavior.

Figure II



Figures II. Fig. II shows the contribution of each demand factor to total price change (measured in percentage points change with respect to the previous year). They were calculated by replacing the estimated values of Z_t^i, A_t,ω, and η ,$i \in$ {food,feed,ethanol,exports} , into each term in equation (2). Each term in equation (2) is represented by a shade of grey for a given year.

Figure IV

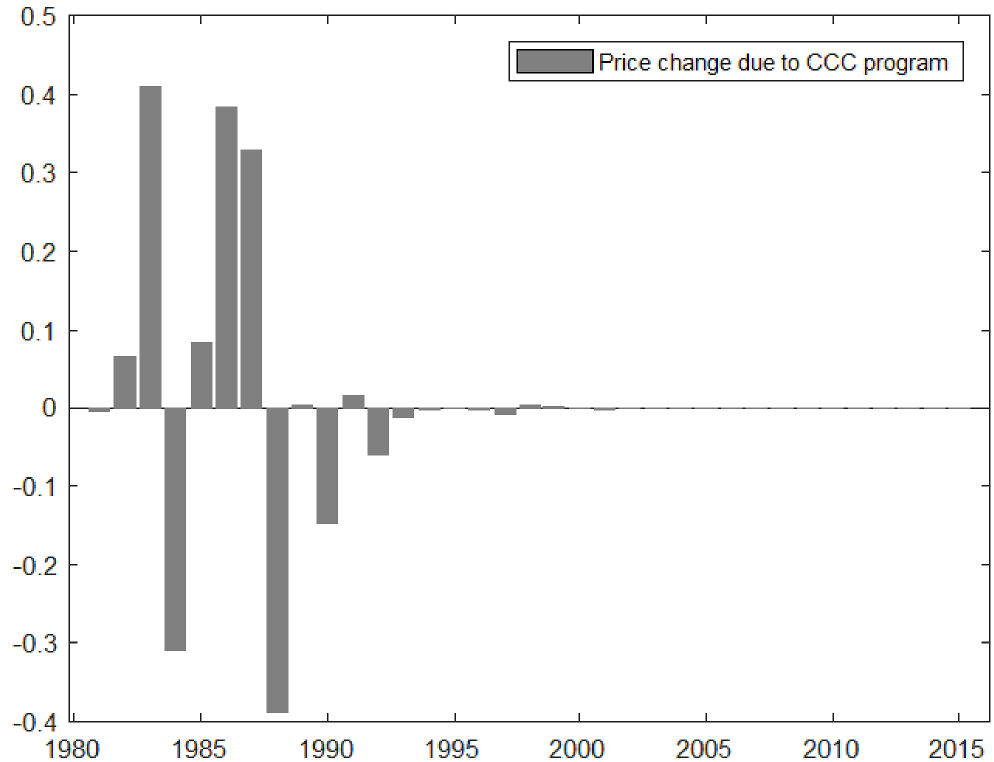


Figure IV. Fig. IV shows the contribution of total price change due to government CCC purchases. It was calculated by estimating the change in prices not explained by equation (2), i.e., the residual between explained price changes and observed price changes. That residual was then multiplied by the proportion of corn inventories held by the government under the CCC program.

3 High Frequency case

The previous model was designed to analyze long-run variation in prices. Our main finding is that there is little evidence of non-fundamental shocks. One may argue that this finding is expected, since, as explained previously, a systematic deviation from our implied fundamental prices would mean either hitting the zero lower bound for inventories or increasing them to infinity, as in Knittel and Pyndick (2016). However, at a higher frequency there could be space for short-term shocks that are not necessarily related to supply and demand fundamentals. Therefore, though I continue to assume isoelastic functional forms for supply and demand, I add speculators as a new type of agent. The justification for this is that given that production of corn occurs only during fall quarter , buying and selling storage inevitably becomes a relevant factor in determining market prices throughout the whole year. That is, give that production only occurs during on quarter, it is unrealistic to ignore inventory changes as a relevant factor. Since in a decentralized framework these inventory decisions are profit-seeking, I rationalize this by introducing a representative speculator: an agent who makes decisions based on expected profits realized by *buying low* and *selling high*. The objective in the next section is to model short-run quarter equilibria and quantify the impact of both current and future shocks (the latter through information shocks) on total variation, similar to what was done in the previous section.

3.1 Model

Time has two dimensions: year and quarter. Notation wise, t represents time change from quarter to quarter, q indicates a quarter-specific notation, and y represents marketing (not calendar) year. A marketing year starts with the harvest and ends exactly before the next one. For simplicity, I will consider demand as a whole (that is, I will not differentiate by use). Just as in the previous model, demand and supply are isoelastic. This seems reasonable, given that the demand elasticities found in the previous section were not so different across the various sources of demand.

$$D_{y,t}(p_{y,t}) = Z_{y,t} p_{y,t}^{-\rho}.$$

$$S_{y,t}(p_{y,t}) = A_{y,t} p_{y,t}^{\eta}.$$

where $Z_{y,t}, A_{y,t}$ are exogenous random variables representing demand and supply fundamentals and have the following dynamics:

$$A_{y,t} = A_y = \rho_A A_{y-1} + \varepsilon_y^A \text{ if } t = 1, A_{y,t} = 0, \text{ for } t \in \{2,3,4\}$$

$$Z_{y,t} = Z_y + \varepsilon_t^Z. \qquad Z_y = \rho_z Z_{y-1} + \varepsilon_y^Z.$$

Shocks are all IID and have the following distributions:

Annual shocks

$$\varepsilon_y^i \sim \log N(0, \sigma_y^i) \qquad \text{for } i \in \{Z, A\}\,.$$

Quarter-specific shocks

$$\varepsilon_t^Z \sim \log N(0, \sigma_q^z),\ \ q \text{ stands for each quarter,1 to 4)}$$

.

I summarize all shocks into vector $\varphi_{y,t} = \left\{\varepsilon_y^Z, \varepsilon_y^A, \varepsilon_t^Z\right\}$

In other words, demand shocks have a yearly component that follows an AR(1) process, plus a quarter-specific noise. Production only comes during one quarter (harvest season); therefore, it is equivalent to define supply shocks as annual or quarterly. In addition to farmers and consumers, speculators, participate in the market by buying and selling stored goods. They are profit-maximizing agents who face the following problem:

$$V_t(x_t, \bar{x}_q) = \qquad \max_{u_t, x_{t+1}} p_t u_t - f(x_t, \bar{x}_q, \delta) + \beta E_t \left[V_{t+1}(x_{t+1})\right].$$

$$\text{s.t.} \qquad x_{t+1} = x_t - u_t$$

$$f(x_t, \bar{x}_q, \delta) = \left\{ \begin{smallmatrix} \delta(x_t - \bar{x}_q)^2 & \text{if} \quad t \in \{1,2,3\} \\ \delta^{II}(x_t - \bar{x}_q)^2 & \text{if} \quad t=4 \end{smallmatrix} \right\} \qquad \delta^{II} > \delta$$

The function $f(x_t, \bar{x}_q, \delta)$ represents the cost of deviation from a certain quarterly optimum $\bar{x}_q$, that has a specific value for each quarter and is known, i.e. it is neither random nor uncertain. Deviating from this optimum implies some cost δ, δ^{II}. The cost in the last quarter is different because usually inventories are managed at low levels during that time of the year. Therefore, given that agents are closer to hitting the zero lower bound, it is reasonable to assume that cost of deviating from optimum are higher. Intuitively, for some industries, having low inventories could imply a huge risk premium: Since they require a minimum level of inventories to keep their machinery running, stopping them (by reaching zero inventories) would imply high costs. As an example, say that annual harvest usually has size 1, and quarter optimums are 0.8,0.6, 0.4, and

0.2, respectively. Having inventory levels below or above the first three does not imply the same costs as having them below 0.2, since in this last case companies would be close to hitting the zero lower bound.

The speculator must decide one period in advance by how much stocks will differ from quarter-specific values.

First-order conditions and envelope conditions are:

$$-p_t + \beta E\left[V'_{t+1}(x_{t+1}, \bar{x}_{q+1})\right] = 0$$

$$-2\delta_{t+1}\left(x_t - \bar{x}_q\right) + \beta E_t\left[V'_{t+1}(x_{t+1}, \bar{x}_{q+1})\right] = \frac{\partial V_t(.)}{\partial x_t} \qquad \delta_{t+1} \in \left\{\delta, \delta^{II}\right\}.$$

Combing both, we get

$$p_t = \beta E\left[p_{t+1}\right] - 2\delta_{t+1}\left(x_{t+1} - \bar{x}_{q+1}\right),$$

Policy function then results in:

$$\begin{aligned} X_{t+1} &= \frac{\beta E\left[p_{t+1}\right] - p_t}{2\beta\delta_{t+1}} + \bar{x}_{q_{t+1}}. \\ \text{with } \delta_{t+1} &\in \left\{\delta, \delta^{II}\right\} \end{aligned} \tag{4}$$

Consumers, farmers, and speculators must then meet at the market. Therefore, the market-clearing equation is given by:

$$X_{y,t+1} = A_y p_{t,y}^{\eta} - Z_{t,y} p_{t,y}^{-\rho} + X_{t,y} \qquad \text{if } t = 1. \tag{5}$$

$$X_{y,t+1} = -Z_{t,y} p_{t,y}^{-\rho} + X_{t,y} \qquad \text{if } t = 2, 3, 4. \tag{6}$$

As a result, given that parameters and market-clearing conditions are quarter specific, equilibrium prices will also be quarter specific. That is, their values will vary from quarter to quarter, even in steady state.

3.2 Information structure

Information is key to forming expectations for speculators, and therefore the information structure is a relevant issue in the model. I construct it to mimic reality as closely as possible.

Production shocks and annual demand shocks are revealed the quarter after they take place, so agents do not know at present moments what supply and demand are. They use both public and private signals to form expectations regarding these. Private signals are formed each quarter. Public signals, on the other hand, are realized only in quarters 1 and 4. I design public signals this way in order to map USDA reports; hence, I must mimic the same timing schedule. Private signals are private in the sense that they are developed by the private sector. They are observable to speculators, but not to the econometrician.

Information regarding demand and supply fundamentals is only revealed after each quarter is finished. During each quarter, therefore the agent does not know with precision what is driving price changes. As an example, in quarter 1, yield, annual demand, and quarter-specific demand shocks are realized, but the agent only knows them after the quarter is over. Therefore, in quarter 2 the annual component of demand is known, but not the quarter-specific one for quarter 2, and so on.

Uncertainty about current shocks takes place only in quarter 1, since in quarters 2 ,3 and 4 they only speculate with respect to next years fundamentals (which by definition have not taken place yet). So, in quarter 1, speculators can observe the current price, but since the current price will depend on the realization of three shocks (annual supply shock, annual demand shock and quarter 1 specific demand shock), they cannot identify or map back to fundamentals (there is no 1 to 1 mapping possible). The assumption used here is that if there is any information that can be captured from this it is negligible and for simplicity will be disregarded.

Agents have access to public reports that are issued in the first and last quarters of the marketing year. The first-quarter report predicts current annual demand and yield and the latter predicts the upcoming year's amounts. Private forecasts regarding next year's yield and demand, on the other hand, are made available every quarter. To summarize, public reports predict current yield (if quarter 1), upcoming yield (if quarter 4), current annual component of demand (if quarter 1), and upcoming demand annual shock (if quarter 4).

To illustrate this, consider the case of a commodities trading company. During the first quarter, it receives the USDA report and the one it privately issues regarding current demand and supply conditions. Once the first quarter has passed, the company observes what happened in quarter 1, and its *research department* continues to develop forecasts for the upcoming year. In contrast the USDA will only make its own forecast once quarter 4 is reached, and so on.

Table II

Public Signaling Timing

Quarter	Harvest	Annual Demand Shock	Public Demand Signal	Public Supply Signal	Productivity Revealing	Demand Revealing
	ε_y^A	ε_y^Z	$\boldsymbol{\theta}_t^{ZPu}$	$\boldsymbol{\theta}_t^{APu}$		
1	X	X	X	X	-	-
2	-	-	-	-	X	X
3	-	-	-	-	-	-
4	-	-	X	X	-	-

Formally, the speculator receives public and private signals at each period t :

In the first quarter (or "harvest" quarter) $t = 1$, signals contain the following structure:

$$\theta_{y,t}^{Apu} = \overbrace{\varepsilon_y^A}^{\text{shock being predicted}} + \overbrace{\eta_t^{pu,A}}^{\text{noise}} \qquad \text{public Yield signal}$$

$$\theta_{y,t}^{Zpu} = \varepsilon_y^Z + \eta_t^{pu,Z} \qquad \text{public Demand signal}$$

$$\theta_{y,t}^{Apr} = \varepsilon_y^A + \eta_t^{pr,A} \qquad \text{private Yield signal}$$

$$\theta_{y,t}^{Zpr} = \varepsilon_y^Z + \eta_t^{pr,Z} \qquad \text{private Demand signal}$$

In non-harvest quarters $t = 2, 3, 4$ signals are:

$$\theta_{y,t}^{Apu} = \overbrace{\varepsilon_{y+1}^A}^{\text{shock being predicted}} + \overbrace{\eta_t^{pu,A}}^{\text{noise}} \qquad \text{public Yield signal}$$

$$\theta_{y,t}^{Zpu} = \varepsilon_{y+1}^Z + \eta_t^{pu,Z} \qquad \text{public Demand signal}$$

$$\theta_{y,t}^{Apr} = \varepsilon_{y+1}^A + \eta_t^{pr,A} \qquad \text{private Yield signal}$$

$$\theta_{y,t}^{Zpr} = \varepsilon_{y+1}^Z + \eta_t^{pr,Z} \qquad \text{private Demand signal}$$

where $\eta_t^{pu,A}, \eta_t^{pu,Z}, \eta_t^{pr,A}, \eta_t^{pr,Z}$ are IID noise components and have a normal distribution with standard deviations (or inverse precision) $\sigma_A^{\theta pu}, \sigma_Z^{\theta pu}, \sigma_A^{\theta pr}, \sigma_Z^{\theta pr}$.

I summarize all signals into vector $\theta_{y,t} = \left\{\theta_{y,t}^{Apr}, \theta_{y,t}^{Zpr}, \theta_{y,t}^{Apu}, \theta_{y,t}^{Zpu}\right\}$

The predicted variable will vary depending on the quarter. In harvest quarters, signals predict the annual demand shock and annual supply shock, which are taking place currently. In non-harvest quarters, signal predict next year's annual demand shock and annual supply shock. This can be seen more clearly in Table III. Therefore, each quarter there will be several sources of shocks: annual demand shocks, annual productivity shocks, quarter-specific demand shocks, and public and private information shocks regarding both demand and supply.

Table III

	Predicted variable			
	Harvest Quarter	Non-Harvest Quarters		
Signal	Quarter 1	Quarter 2	Quarter 3	Quarter 4
$\theta_{y,t}^{Apu}$	ε_y^A	-	-	ε_{y+1}^A
$\theta_{y,t}^{Zpu}$	ε_y^Z	-	-	ε_{y+1}^Z
$\theta_{y,t}^{Apr}$	ε_y^A	ε_{y+1}^A	ε_{y+1}^A	ε_{y+1}^A
$\theta_{y,t}^{Zpr}$	ε_y^Z	ε_{y+1}^Z	ε_{y+1}^Z	ε_{y+1}^Z

Table III. This table shows what variable is being predicted by each signal in each quarter. In the harvest quarter, signals intend to predict *current* shocks, since these not inmediately observable. In the rest of the year signals intend to predict *next year's* shocks.

Once a signal is received, agents use it to update their previous beliefs. That is, they use past private (public) signals and combine them with their latest private (public) signal, constructing a posterior private (public) signal. For example, in quarter 2, the private demand signal the agent receives is the first one that forecasts next year's fundamentals. In this case she has a flat prior, and therefore her posterior will be identical to the signal received. In the third quarter, however, the agent has a previous private demand signal θ_2^{Zpr} and receives θ_3^{Zpr}. Assuming both signals have equal precision, she weights them optimally by assigning equal weight to both (recall that the noise process is IID). Therefore, in the third quarter, posterior signal $\check{\theta}_3^{Zpr}$ will be obtained by following a Bayesian updating process with normally distributed noise:

$$\check{\theta}^{Zpr}_{3,y} = 0.5\theta^{Zpr}_{2,y} + 0.5\theta^{Zpr}_{3,y} \qquad \text{with standard deviation} \quad \check{\sigma}^{\theta pr}_{3,y,Z} = \frac{1}{\sqrt{2}}\sigma^{\theta pr}_{Z}.$$

In the fourth quarter, she receives another private demand signal, which she again uses to update:

$$\check{\theta}^{Zpr}_{4,y} = \frac{1}{3}\theta^{Zpr}_{2,y} + \frac{1}{3}\theta^{Zpr}_{3,y} + \frac{1}{3}\theta^{Zpr}_{4,y} \qquad \text{with standard deviation} \quad \check{\sigma}^{\theta pr}_{4,y,Z} = \frac{1}{\sqrt{3}}\sigma^{\theta pr}_{Z}.$$

In the first quarter the agent receives the last private demand signal

$$\check{\theta}^{Zpr}_{1,y+1} = \frac{1}{4}\theta^{Zpr}_{2,y} + \frac{1}{4}\theta^{Zpr}_{3,y} + \frac{1}{4}\theta^{Zpr}_{4,y} + \frac{1}{4}\theta^{Zpr}_{1,y+1} \qquad \text{with standard deviation} \quad \check{\sigma}^{\theta pr}_{1,y+1,Z} = \frac{1}{\sqrt{4}}\sigma^{\theta pr}_{Z}.$$

For private signals regarding supply, constructions are identical. In the case of public signals, calculations are also identical, but since they only take place for two quarters (4 and 1), they have final precision $\sigma^{\theta pu}_{A}$, $\frac{1}{\sqrt{2}}\sigma^{\theta pu}_{A}$ for supply and $\sigma^{\theta pu}_{Z}$, $\frac{1}{\sqrt{2}}\sigma^{\theta pu}_{Z}$ for demand, respectively, for each quarter.

Agents therefore have, each quarter, final public demand and supply signals, and final private demand and supply signals. To rationally summarize this information, these final signals are weighted as a function of their posterior precision. Formally, in quarters 1 and 4, they have posterior public signals $\theta^{Apu}_{y,t}, \theta^{Zpu}_{y,t}$ with a noise process that also follows a normal distribution, with precision $\check{\sigma}^{pu^{-1}}_{A}, \check{\sigma}^{pu^{-1}}_{Z}$ calculated above. They then update their beliefs according to Bayes' rule, that is, they construct a single terminal signal for demand and a single terminal signal for supply, based on a precision-weighted average of both posterior private and public signals:

$$\theta^{A}_{y,t} = \gamma_{t,A}\check{\theta}^{Apu}_{y,t} + \left(1-\gamma_{t,A}\right)\check{\theta}^{Apr}_{y,t} \qquad \check{\theta}^{Apu}_{y,t}, \check{\theta}^{Apr}_{y,t} \text{ with precisions } \check{\sigma}^{pu^{-1}}_{t,y,A}, \check{\sigma}^{pr^{-1}}_{t,y,A}.$$

$$\theta^{Z}_{y,t} = \gamma_{t,Z}\check{\theta}^{Zpu}_{y,t}, + \left(1-\gamma_{t,Z}\right)\check{\theta}^{Zpr}_{y,t} \qquad \check{\theta}^{Zpu}_{y,t}, \check{\theta}^{Zpr}_{y,t} \text{ with precisions } \check{\sigma}^{pu^{-1}}_{t,y,Z}, \check{\sigma}^{pr^{-1}}_{t,y,Z}.$$

$$\gamma_{4,A} = \frac{\check{\sigma}^{\theta pu^{-1}}_{4,y,A}}{\check{\sigma}^{\theta pu^{-1}}_{4,y,A} + \check{\sigma}^{\theta pr^{-1}}_{4,y,A}} \qquad \gamma_{4,Z} = \frac{\check{\sigma}^{\theta pu^{-1}}_{4,y,Z}}{\check{\sigma}^{\theta pu^{-1}}_{4,y,Z} + \check{\sigma}^{\theta pr^{-1}}_{4,y,Z}} \qquad \text{for quarter 4,}$$

$$\gamma_{1,A} = \frac{\check{\sigma}^{\theta pu^{-1}}_{1,yA}}{\check{\sigma}^{\theta pu^{-1}}_{1,y,A} + \check{\sigma}^{\theta pr^{-1}}_{1,y,A}} \qquad \gamma_{1,Z} = \frac{\check{\sigma}^{\theta pu^{-1}}_{1,y,Z}}{\check{\sigma}^{\theta pu^{-1}}_{1,y,Z} + \check{\sigma}^{\theta pr^{-1}}_{1,y,Z}} \qquad \text{for quarter 1,}$$

Parameters $\gamma_{t,A}, \gamma_{t,Z}$ $t \in \{4, 1\}$ indicate the weight agents put in each variable on quarters 4 and 1. Such weights depend on the relative precision of each final signal.

Definition 1 *Speculators do not observe current A and Z; instead they form expectations based on private and public signals regarding demand and supply. Hence, their information set $\Theta_{y,t}$ can be formally defined as:*

$$\text{For quarter } t = 1\text{: } \Theta_{y,t} = \left\{ \underbrace{A_{y-1}}_{\text{annual productivity}}, \underbrace{Z_{y-1}}_{\text{demand level}}, \underbrace{\varepsilon^{z}_{y-1,4}}_{\text{previous quarter demand shock}}, \underbrace{X_{y,1}}_{\text{inventory}}, \underbrace{\theta^{A}_{y,1}, \theta^{Z}_{y,1}}_{\text{signals}}, \Theta_{y-1,4} \right\}.$$

$$\text{For quarters } t = 2,3,4\text{: } \Theta_{y,t} = \left\{ \begin{array}{c} \underbrace{A_{y}}_{\text{annual productivity}}, \underbrace{Z_{y}}_{\text{demand level}}, \underbrace{\varepsilon^{z}_{y,t-1}}_{\text{previous quarter demand shock}}, \\ \underbrace{X_{y,t}}_{\text{inventory}}, \underbrace{\theta^{A}_{y,t}, \theta^{Z}_{y,t}}_{\text{signals}}, \Theta_{y,t-1} \end{array} \right\}.$$

Therefore, expected prices can be defined as:

$$E(p_{y,t}) = E(p_{y,t}|\Theta_{y,t}).$$

$\Theta_{y,t} \in \Omega$, where Ω is the set that contains all information and demand and supply shocks, and therefore contains $\Theta_{y,t}$ $\forall$ $\{y, t.\}$ Note that by construction, $\theta_{y,t} \in \Theta_{y,t}, \varphi_{y,t-1} \in \Theta_{y,t}$ $\forall$ $\{y, t.\}$ and Information sets are a combination of current signals and past shocks, that come from their own distributions. Therefore Ω is a measurable space and $\left\{\Theta_{y,t}, \varphi_{y,t-1}\right\} \in \{\Omega, \mathcal{F}, P\}$

The solution to the equilibrium system of equations I will be looking for will have the following functional—nonlinear—form,given that although speculators have a linear-quadratic objective, the rest of the supply and demand functions are isoelastic:

$$p_{y,t} = p_{y,t}(\Theta_{y,t}, \varphi_{y,t}) = p(\underbrace{A_{y}}_{\text{productivity}}, \underbrace{Z_{y}}_{\text{demand level}}, \underbrace{\varepsilon^{z}_{y,t}}_{\text{quarter shock}}, \underbrace{X_{y,t}}_{\text{Inventory}}, \underbrace{\theta^{A}_{y,t}, \theta^{Z}_{y,t}}_{\text{signals}}, A_{y-1}, Z_{y-1}). \tag{7}$$

Definition 2 *An equilibrium consists of prices $p_{y,t}(\Theta_{y,t}, \varphi_{y,t})$ such that, given shocks $\{\varphi_{y,t}, \theta_{y,t}\}$, information set $\Theta_{y,t}$, and initial conditions $X_{0,1}, A_{-1}, Z_{-1}$, markets clear for all quarters and the speculator's FOC is satisfied for all $\{y, t\}$*

Prices will depend not only on current level of productivity and demand, but also on the level of initial inventories and available signals. Past productivity and demand can influence this by providing information about future realizations of demand and supply, since they follow an AR(1) process. Given this solution, two main issues remain to be addressed: First and most important, the solution is nonlinear, which makes it impossible to find an explicit solution. Second, private signals are non-observable in the data. Therefor, I will solve the model by proceeding through the following steps:

Step 1

Linearize the system around steady-state values. I normalize these to 1, such that index points can be interpreted as being close to percentage changes.

Step 2

Calibrate parameters using USDA corn market data for 1975-2016.

Step 3

Estimate seasonality and simulate the model by feeding estimated shocks as inputs, the compare model prices versus data.

Step 4

Estimate private signals using a Kalman Filter.

Step 5

Estimate the contribution of each factor to final prices and measure the residual (unexplained fraction of price changes).

3.2.1 *Step 1*

I proceed to linearize the model around steady state. Variables are redefined in the following way:

For any variable y, $\hat{y} = y - \bar{y}$ where $\bar{y}$ is the steady-state value of the respective variable and $\hat{y}$ is the linear deviation from it.

Therefore, we must solve for steady-state values. Steady-state estimates are obtained in the following way:

First, I define a harvest size of one:

$$\bar{A} = 1.$$

Then I normalize demand for each quarter, such that they all add up to the harvest

$$\sum_i^4 Z_i = 1.$$

I estimate quarter-specific seasonality $\bar{Z}_i$ for demand in the data and normalize it such that they all add up to the harvest size:

$$\bar{Z}_i' = \frac{\bar{Z}_i}{\sum_i^4 \bar{Z}_i}.$$

Therefore, we will have $\sum_i^4 \bar{Z}_i' = \bar{A} = 1.$

For inventories, in steady state they should not deviate from $\bar{x}_q$ since the values for these are picked such that sticking to the target is always optimum. If speculator's match their target $\bar{x}_q$, the marginal inventory cost is zero These are calibrated by taking average inventory levels (relative to harvest size) for each quarter throughout the whole sample. Therefore, the policy function equation from the speculator's problem will give us the evolution of prices throughout each quarter in steady state:

$$\bar{X}_{t+1} - \bar{x}_{q+1} = 0 \qquad \Rightarrow \qquad \bar{p}_4 = \frac{1}{\beta}\bar{p}_3 = \frac{1}{\beta^2}\bar{p}_2 = \frac{1}{\beta^3}\bar{p}_1.$$

Therefore, we only need the value of prices in steady state for one quarter to determine the rest. Given values $\bar{A}, \bar{Z}_i', \bar{X}_i$, $i = 1, 2, 3, 4$ and the price evolution path, I can introduce these values into steady-state market clearing equations to obtain $\bar{p}_1$ numerically:

$\bar{X}_2 = \bar{A}p_1^{\eta} - \bar{Z}_1'\bar{p}_1^{-\rho} + \bar{X}_1.$

$\bar{X}_3 = -\bar{Z}_2'\bar{p}_2^{-\rho} + \bar{X}_2.$

$\bar{X}_4 = -\bar{Z}_3'\bar{p}_3^{-\rho} + \bar{X}_3.$

$\bar{X}_1 = -\bar{Z}_4'\bar{p}_4^{-\rho} + \bar{X}_4.$

Once steady-state values are calculated, I proceed to linearize previous equations (4) and (5).

Linearizing the market-clearing equation results in:

$$\hat{X}_{t+1,y} = \eta \bar{A}_y \bar{p}_{t,y}^{\eta-1}\hat{p}_{t,y} + \bar{p}_t^{\eta}\hat{A}_y - \bar{p}_{t,y}^{-\rho}\hat{Z}_{t,y} + \rho\bar{Z}\bar{p}_{t,y}^{-\rho-1}\hat{p}_{t,y} + \hat{X}_{t,y}. \tag{8}$$

Speculator's policy function will then become:

$X_{t+1} = \frac{\beta E[p_{t+1}] - p_t}{2\beta\delta^{t+1}}. \qquad \Rightarrow \qquad 2\beta\delta^{t+1}X_{t+1} = \beta E\left[p_{t+1}\right] - p_t$

$$\begin{aligned} 2\beta\delta^{t+1}\hat{X}_{t+1} &= \beta E\left[\hat{p}_{t+1}'\right] - \hat{p}_t \\ \text{with } \delta^{t+1} &\in \left\{\delta, \delta^{II}\right\}. \end{aligned} \tag{9}$$

Therefore, variables will reflect deviation from steady-state values in levels. It is important to point out that steady-state values are quarter specific. Equations (8) and (9) are linear in prices and inventories. Therefore, the price solution equation will have also a linear solution. Since both equations depend on parameters that vary in each quarter, the price solution equation will also vary depending on the quarter.

Conclusion 3 *As a result, the linearized new state space will have four price solution equations, one for each quarter:*

$$\hat{p}_1 = \alpha_1^I \hat{A}_1 + \alpha_2^I \varepsilon_1^Z + \alpha_3^I \hat{X}_1 + \alpha_4^I \theta_1^A + \alpha_5^I \theta_1^Z + \alpha_6^I \hat{Z}_{y-1} + \alpha_7^I \hat{A}_0 + \alpha_8^I \hat{Z}_y.$$

$$\hat{p}_2 = \alpha_1^{II} \hat{A}_1 + \alpha_2^{II} \varepsilon_1^Z + \alpha_3^{II} \hat{X}_1 + \alpha_4^{II} \theta_1^A + \alpha_5^{II} \theta_1^Z + \alpha_6^{II} \hat{Z}_{y-1} + \alpha_7^{II} \hat{A}_0 + \alpha_8^{II} \hat{Z}_y.$$

$$\hat{p}_3 = \alpha_1^{III} \hat{A}_1 + \alpha_2^{III} \varepsilon_1^Z + \alpha_3^{III} \hat{X}_1 + \alpha_4^{III} \theta_1^A + \alpha_5^{III} \theta_1^Z + \alpha_6^{III} \hat{Z}_{y-1} + \alpha_7^{III} \hat{A}_0 + \alpha_8^{III} \hat{Z}_y.$$

$$\hat{p}_4 = \alpha_1^{IV} \hat{A}_1 + \alpha_2^{IV} \varepsilon_1^Z + \alpha_3^{IV} \hat{X}_1 + \alpha_4^{IV} \theta_1^A + \alpha_5^{IV} \theta_1^Z + \alpha_6^{IV} \hat{Z}_{y-1} + \alpha_7^{IV} \hat{A}_0 + \alpha_8^{IV} \hat{Z}_y.$$

$$\theta_i^A = [\theta_i^{Apr} \qquad \theta_i^{Apu}] \qquad \theta_i^Z = [\theta_i^{Zpr} \qquad \theta_i^{Zpu}].$$

where $\alpha_i^I, \alpha_i^{II}$, α_i^{III}, α_i^{IV} are price policy function parameters with $i = 1, .., 8$ for each state variable and $q = I, .., IV$ for each quarter. Recall that in steady state $\theta_i = 0$ for all signals θ, both private and public.

It is important to remember that the sole role of "last year's" production and demand is to inform about future production and demand. Therefore, if persistence parameters ρ_A and ρ_Z are equal to zero, α_7^J and α_8^J should also be zero for all J.

3.2.2 *Step 2*

I obtain USDA data series for yield per acre, production, demand (local and exports), prices, and stocks. I proceed first by detrending and deseasonalizing them, then normalize their mean so that index points can be read as percentage points. To obtain the exogenous demand component, I instrument prices with yield per acre to obtain demand elasticity through a typical IV analysis. I use residuals as estimates for $Z_{y,t}$, just as in the long-run-trend model. I then estimate the annual component of the residuals to obtain Z_y and Z_t series. For supply elasticity η, again I use Roberts & Shlenker's (2013) results. As a robustness check, other values for η were used with very similar results.

When it comes to reports, I digitalize USDA forecasts for each marketing year, from 1975 to 2016. These forecasts are for demand and yield per acre. To obtain θ_t^{Zpu} forecasts from demand forecasts, I regress these last ones against forecasted yields, then use the residual as a demand shock estimate. I do this because given that a simple demand forecast contains also an endogenous forecast, since a higher yield would endogenously result in lower prices and higher consumption, I eliminate this endogenous component of the demand forecast by regressing it against the yield forecast.

In this section I calibrate the parameters to simulate and show quantitative results. Given the data available (described above), I do the following. First, I calibrate the parameters necessary to calculate α_i^q which are $\beta, \delta^I, \delta^{II}, \bar{A}, \bar{Z}_y, \bar{Z}_q, \omega$, and η. Demand and supply elasticities are estimated as indicated in the previous step. The persistence of supply and demand fundamentals are calibrated by estimating AR(1) processes, and volatility is obtained by taking the residual's standard deviations.

Storage cost parameters δ^I, δ^{II} are estimated by choosing values that minimize residuals. That is, the previous set of equations will give theoretical prices as the result for each quarter. The difference between these and observed prices are the residual. Parameters δ^I, δ^{II} were calibrated such that the total sum—in absolute values—of the residuals is minimized.

The next step is to calibrate public and private signaling parameters. Regarding public signals, it is important to say that these appear in specific days, which are publicly known in advance. That is, at a certain time, information becomes publicly known in a report and markets react almost instantaneously Price changes for those "announcement" days can be seen, and even though private information is always present in markets, it is reasonable that by constraining price changes to the specific announcement day isolates public signal's effect on prices.

The precision of public forecasts is estimated by comparing USDA demand and supply forecasts versus ex post realizations. The most non-straightforward task, however, is to obtain a value for the precision of private signals. I proceed to identify them as follows. First, I calculate α_4^I and α_5^I by using the method of undetermined coefficients (I use this to obtain all parameters α_i^q). Second, I obtain the release dates for public signals (USDA forecast reports) and the same-day price reactions. I then regress same-day price changes against the changes in forecasted values in the released reports—with respect to the previous report—for that day. This allows me to estimate the effect of public signals on price changes. I know that the parameter that relates price changes to public supply signaling in the model, $\alpha_4\gamma_{1,A}$, should match this previous price reaction estimation. Once I know the value for $\alpha_4\gamma_{1,A}$, and given that I also know α_4, I can obtain the value of the weight of public signals relative to private ones, $\gamma_{1,A}$ Since $\gamma_{1,A} = \frac{\check{\sigma}_{1,yA}^{\theta pu^{-1}}}{\check{\sigma}_{1,y,A}^{\theta pu^{-1}} + \check{\sigma}_{1,y,A}^{\theta pr^{-1}}}$ and public signal precision, $\check{\sigma}_A^{\theta pu}$, is known, I can identify private signal precision, $\check{\sigma}_A^{\theta pr}$. The procedure to identify

$\breve{\sigma}_Z^{\theta pr}$ is identical.

Calibration results for US corn markets can be seen in Table IV.A:

Table IV.A

Description	Variable	Value	Source
Discount factor	β	0.995	Literature
Demand elasticity	ω	0.18	IV
Supply elasticity	η	0.15	Literature
Demand persistence	ρ_z	0.64	AR(1) Z_t
Supply persistence	ρ_A	0	AR(1) A_t
Steady State productivity	$\bar{A}$	1	Normalization
Steady State demand level	$\bar{Z}_y$	0.25	Normalization
Inventory adjustment cost quarters I,II,III	δ^I	0.0204	Model calibration
Inventory adjustment cost quarter IV	δ^{II}	0.046	Model calibration
Yield volatility	σ_A	0.098	Stand Dev A_t
Demand volatility	σ_Z	0.018	Stand Dev Z_t
Public signal A st dev	$\sigma_A^{\theta pu}$	0.06	USDA Yield forecast
Public signal Z st dev	$\sigma_Z^{\theta pu}$	0.048	USDA Demand forecast

3.2.3 *Step 3*

Once I have determined the values for the parameters, I simulate the model by feeding in shocks obtained from data. That is, I normalize, detrend, and deseasonalize the data and use the residuals as supply, demand, information, and inventory state variables. Since I don't observe private signals, for now I will assume them to be zero (I will later proceed to estimate them). Simulation results can bee seen in Figure VIII.

Figure VIII

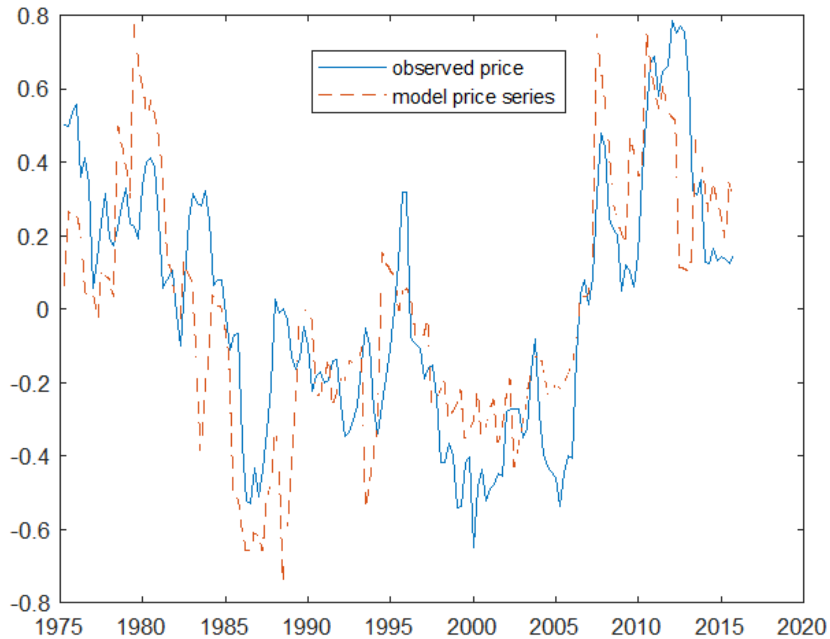


Figure VIII. Lines show the evolution of observed corn prices versus model simulated ones. Units are deviations from steady state in percentage points.

Prices Figure VIII shows the results given by the simulation versus those in the data. At first glance, we can say the model generates a price series that is in line with the one observed in the data. However, it is clear that for some years there are discrepancies between model-generated and observed prices. This can be better seen in Figure IX, which shows residuals measured as differential between observed and effective prices. On average, around 36% of price changes cannot be explained within the model. However, this is not uniform throughout time. Indeed, there are quarters, such as 1985-1987, in which non-modeled factors explain almost 60% of price changes, and others in which that drops to less than 10%. There are several reasons for this. From a theoretical point of view, the model is a linear approximation to equilibrium values. Hence, second-order effects might be bigger than expected. From a practical perspective, there could be non-profit-maximizing institutions in place—for example, government policy. Agriculture is an industry subject of several policies, with tax breaks, subsidies, inventory, and price policies in all world markets, with the US being no exception. It is important to say, however, that unexplained factors do not seem to have increased in the past 15 years.

Figure IX

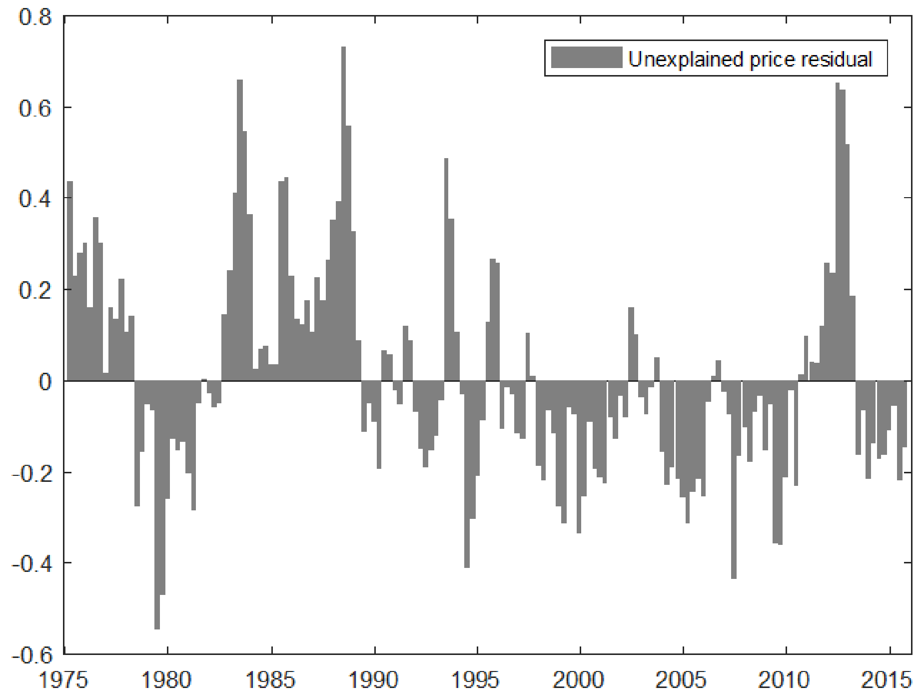


Figure IX. Bars show the percentage points of price changes that cannot be explained by factors in the model. For some periods of time, prices can differ almost 60% from these based on their fundamentals.

3.2.4 *Step 4*

A key innovation of the model is the incorporation of both private and public signals into agents' decisions. Agents use these signals to predict the annual component of demand and yields. They have both private and public sources, and weight them proportional to their precision. I use USDA reports released monthly that predict future demand and yield per acre; I consider these to be *public signals* since anyone can access them at no cost. I then estimate private signals through a

Kalman Filter, in which the observables are prices and ex post realized annual demand and yield values.

Formally:

$x_t = F_t x_{t-1} + \varepsilon_t.$ state space model

$y_t = H_t x_t + B_t u_t + \upsilon_t.$ observables

where

$y_t = [p_t \; \varepsilon^A_{t+1} \; \varepsilon^Z_{y+1}].$

$u_t = [A_t \; \varepsilon^Z_t \; X_t \; \theta^{Apu}_t \; \theta^{Zpu}_t \; Z_{y-1} \; Z_y].$

$x_t = [\theta^{Apr}_t \; \theta^{Zpr}_t].$

Σ_υ and Σ_ε are covariance matrices for $\upsilon_t, \varepsilon_t$. I calibrate Σ_υ by measuring the precision of public forecasts and the standard deviation of changes in price that remain unexplained by the model. Values for Σ_ε,are the precision in private forecasts estimated previously. Matrices F_t, H_t, and B_t represent the laws of motion for private signals, the effect of private signals on prices, and the effect of the rest of state variables on prices, respectively. Also, the second and third rows of matrix H_t represent the relation between signals and ex post realizations. That is, they capture the fact that signals forecast ex post realized demand and supply shocks with some error.

Estimations can be seen in Figures X and XI, which show the estimated private signal. Although private and public signals tend to move together, as one would expect, there are several moments in which these differ, which may explain public information is not the only driver of expectations.

Figure X

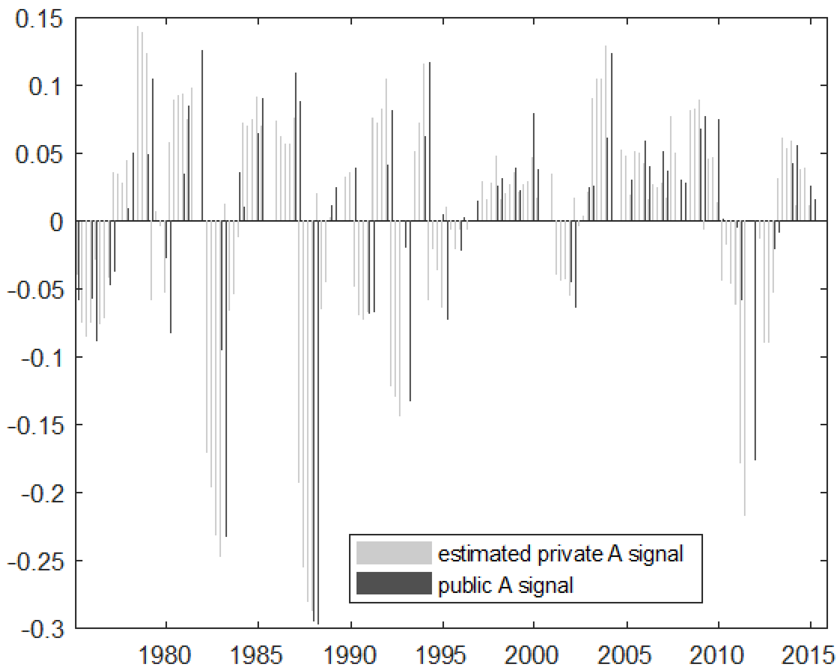
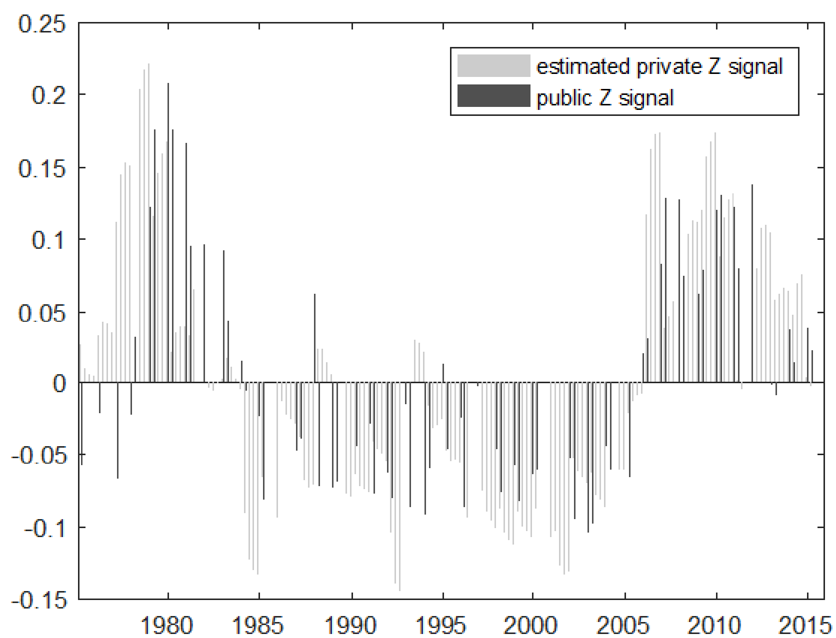


Figure XI



Figures X - XI. Bars show the forecasted demand and supply shock for each quarter by both private and public signals. Fig. X shows supply shock forecasts and Fig. XI shows demand shock forecasts. Darker bars show forecasts from USDA, and light bars show private forecasts estimated using a Kalman filter.

An important step is not only estimating signals, but also to checking if they are relevant. Figures XII and XIII (see Appendix) show the contribution to price change in index points for 1975-1979. In certain quarters, their impact on prices can reach almost 10%. That is, these information shocks are economically significant the moment they occur. When considered over the whole sample time span, their relevance differs based on whether they are public or private. In the first case, overall impact drops to 2.6% (combining demand and supply forecasts), while in the latter it reaches 15%. This is not surprising, since in the model public signals only occur in two quarters each year, whereas private signals take place every quarter. Also, private signals are more volatile, and therefore have a greater impact in total volatility.

3.2.5 *Step 5*

Table IV.B describes the contribution of each factor to price changes, measured by the average absolute value of factor effect over price change (according to model policy parameters). As expected, the main driving factors are demand components (both annual and quarterly) and inventories, since when they are combined we reach almost 50% of price variation. The number associated with yield per acre is particularly small, but this should not surprise since the model divides time into quarters, with production (yield) only in one of four quarters. The supply influence on prices, hence, tends to be absorbed by stocks in the following quarters, since when there is a good or bad yield this later changes the level of stocks during the rest of the marketing year.

Figure XIV

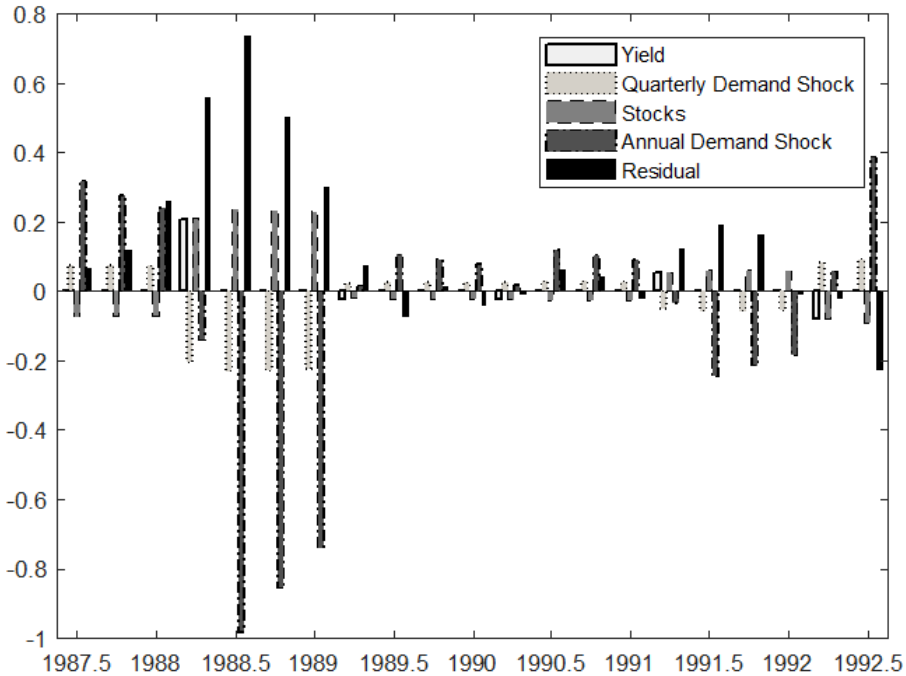


Figure XIV. Decomposition of price changes per factor. Different shades indicate the individual effect of each variable over total price change for that quarter. Units are in percentage points.

Table IV.B

Fraction of Price Change Due to Each Variable*		
Variable	Description	Fraction
θ_t^{Apr}	Supply Private Signal	0.056
θ_t^{Zpr}	Demand Private Signal	0.099
A_t	Productivity	0.018
ε_t^Z	Quarter Demand Shock	0.089
X_t	Inventories	0.088
θ_t^{Apu}	Supply Public Signal	0.009
θ_t^{Zpu}	Demand Public Signal	0.019
Z_{y-1}	Past Demand Level	0.043
A_{t-1}	Past Productivity	0
ε_y^Z	Annual Demand Shock	0.274
Residual		0.305

*in absolute values

4 Conclusion

This paper addresses the issue of agricultural commodity price changes throughout the past decades. In particularly, I take the case of corn for US markets. I derive two analyses: long run and short run. In the first, I develop a simple model with annual variables in which I simulate theoretical prices, such that demand and supply match each year. That is, I estimate the price that would have theoretically cleared the market (no inventory changes), then compare it against the observed prices. The intention was to develop what one could consider a proxy for a non-distorted price and compare to see whether the observed price was too far away from it. Results show that after the 1996 farm bill, the relationship between these two time series is very close. That is, there does not seem to be a major difference between the market-clearing price and the observed price. Before 1996, observed prices clearly differed from theoretical ones. There are several explanations for this, but a major one is that government agriculture policy intervened heavily, mostly through price-sustaining policies. These were instrumented by either government accumulation of stocks (the CCC program) or subsidizing the private sector to do that for them (the FOR program). After 1996, government policy became more market friendly by introducing price floors below equilibrium and limiting the government's budget available for stocks. As a first result, Figure I shows that no big distortions are to blame for price hikes in the 2000s. Moreover, a more detailed analysis shows that the emergence of an energy-related source of demand for corn had a significant impact, and explains an important aspect of price increases in 2005-2010. This is in line with the increasing positive correlation between corn prices and energy ones. An important thing to note is that government mandates could have played a major role in this. Therefore, it could be that we are observing a shift in regulation type from direct purchases to private coercion. Identifying these interventions, quantifying them, and estimating their economic impact are subject's for future research.

The second part of the paper analyzes the short run. I develop a quarterly based model in which agents can only produce at certain quarters, as in most agricultural markets. I then linearize it and analyze second order moments. When focusing on a higher frequency, there is space for speculation to play an important role, since it is possible to make profits by storing for short periods of time and reselling. In other words, it is irrational for inventories to go up *forever*, but it is perfectly possible that they increase for one or two quarters. Therefore, I proceed by modeling speculation and introduce a theoretical innovation: information shocks. Given that storage is by definition related to expectations about future prices, the information available at each moment is key to inventory decision-making. I introduce two sources of information: public and private. Public information comes as a forecast by a government agency (in this case, the USDA). Private information comes from information markets (consulting firms, private research departments, etc.). These reports help agents form expectations about what future prices might be and make decisions accordingly. Given that USDA reports are publicly available, and that they contain quantitative forecasts for demand and supply, I am able to estimate their impact on market prices. I also identify the precision of private signals and estimate a time series for them. Results show that as one would expect, demand and supply shocks account for a large fraction of price changes—as much as 52%. However, this leaves a relevant space for speculation or inventory shocks. Information shocks (which act by inducing agents to buy or sell their stored goods) can explain an additional 18%. A remaining 31% is due to factors not included in the model but that alter storage decisions, as well as second-order effects in demand or supply. That is, in the second section I am able to quantify by how much these unexplained factors can influence observed prices. Finally, another important result is that non-explained sources of price changes, whatever these may be, do not seem to be any more relevant in the 2000s than they were before. This again confirms once more that the idea that nonfundamental factors have increased in importance is not sustainable.

Results show that in the short run, prices can deviate severely from fundamentals, even though in the long run that is not the case. Given that the market analyzed here is well developed in terms of liquidity and transparency, this raises questions regarding how informative prices may be in the short run, with important implications with respect to GDP accounting and theoretical modeling. As in Romer (2016), my results challenge the notion of market prices as a result of a standard profit- and utility-maximizing framework.

In conclusion, the evidence presented here indicates that there is no evidence for the hypothesis that corn markets artificially deviated from fundamentals in the long run, but does allow the possibility for quarterly frequencies. As an innovation, this paper also shows the important role information shocks play in the short run. Unknown sources of variability are quantified, but are yet to be explained. Further research in this respect is necessary.

5 References

Acharya, V.V., Lochster, L.A. (2013), Limits to arbitrage and hedging : Evidence from commodity markets. *Journal of Financial Economics*, vol. 109, issue 2, 441-465.

Albagli, E., Hellwig, C. and Tsyvinksi, A. (2015), A Theory of Asset Prices based on Heterogeneous Information, *revise and resubmit to Review of Economic Studies.*

Basak, S. & Pavlova, A. (2016), A Model of Financialization of Commodities, *Journal of Finance*, Vol. LXXI, No 4.

Caldara, D. ,Cavallo M. & Iacoviello M. (2016), Oil Price Elasticities and Oil Price Fluctuations, *International Finance Discussion Papers* 1173, Board of Governors of the Federal Reserve System (U.S.).

Fattouh, B., Kilian L., and Mahadeva L. (2013), The Role of Speculation in Oil Markets: What Have We Learned So Far?, *The Energy Journal*, vol. 34, issue 3.

Bohl, M.T. & Stephan, P.M. (2012), Does Futures Speculation Destabilize Spot Prices? New Evidence for Commodity Markets, *The Journal of Future Markets*, August, 696-714.

Bosch, D. & Pradkhan, E. (2015), The impact of speculation on precious metals futures markets, *Resources Policy*, v 44, 118-134.

Chari, V.V. & Christiano, L.J. (2017), Financialization in Commodity Markets, *Federal Reserve Bank of Minneapolis*, Staff Report 552.

Chahrour, R. & Kyle, J. (2018), News or Noise? The Missing Link, *American Economic Review*, 108 (7): 1702-36.

Egelkraut, T.M., Garcia, P. , Iriwn, S.H. and Good, D.L. (2003), An evaluation of Crop Forecast Accuracy for Corn and Soybeans: USDA and Private Information Agents, *Journal of Agricultural and Applied Economics*, April 35,1:79-95.

Hansen, L.P. & Sargent, T.J. (1993), Seasonality and approximation errors in rational expectations models, *Journal of Econometrics*, 55-21.

Hussman, J.P. (1992), Market Efficiency and Inefficiency in Rational Expectations Equilibria: Dynamic Effects of Heterogeneous Information and Noise, *Journal of Economic Dynamics and Control*, 16: 655-680.

Iqbal, Z. & Babcock, B.A. (2018), Global Growing Area Elasticities of Key Agricultural Commodities Estimated Using Dynamic Heterogeneous Panel Methods, *Agricultural Economics*, September, Vol 49, Issue 5, 547-668.

Kilian, L. & Murphy, D. (2014), The Role of Inventories and Speculative Trading in the Global Market for Crude Oil, *Journal of Applied Econometrics*, April, 29(7753)

Knittel, C.R. & Pyndick, R. (2016), The Simple Economics of Commodity Price Speculation, *American Economic Journal: Macroeconomics* , 8(2): 85–110.

Mehra, R. (2012), Consumption-Based Asset Pricing Models, *The Annual Review of Financial, Economics* 4:385–409.

Milgrom, P. & Stokey, N. (1982), Information, Trade and Common Knowledge, *Journal of Economic Theory*, vol. 26, issue 1, 17-27.

Pindyck, R. (1990), Inventories and the Short-Run Dynamics of Commodity Prices, *The RAND Journal of Economics*,vol. 25, no. 1, 1994, pp. 141–159.

Roberts, M.J., and Wolfram, S.(2013). "Identifying Supply and Demand Elasticities of Agricultural Commodities: Implications for the US Ethanol Mandate. " *American Economic Review"*, 103 (6): 2265-95.

Romer, D. (2006), Do Firms Maximize? Evidence from Professional Football, *Journal of Political Economy*, 114(2):340-365.

Scheinkman, J. & Xiong, W. (2003), Heterogenous Beliefs, Speculation and Trading in Financial Markets, *Paris-Princeton Lectures in Mathematical Finance 2003, Lecture Notes in Mathematics*, vol 1847, pp 217-250

Sockin, M. & Xiong, W, (2015), Informational Frictions and Commodity Markets, *The Journal of Finance,* October, Vol. LXX, No. 5.

Townsend, R.M. (1983), Forecasting the Forecast of Others, *Journal of Political Economy*, Vol. 91, No 4, pp. 546-588.

Part II

Commodity Futures and Seasonal Correlation

Francisco Arroyo Marioli, UCLA

Although competitive storage theory has proven successful in explaining many patterns for commodity prices, some important features remain unexplained. Particularly, while standard models predict low correlation between future prices with delivery dates before and after the harvest, the data suggests otherwise. To correct this issue, my approach assumes that harvests appear continuously rather than at a single moment. This addition to the standard model allows me to link pre-harvest and post-harvest markets to the same source of supply, hence obtaining the high correlation observed in the data. Empirical evidence also suggests that the assumptions used are realistic. Results are robust to different parameter specifications.

6 Introduction

Commodity markets have several characteristics that define them and make them unique. They are more homogeneous, more transparent and more liquid than other types of goods. Information regarding price and quantity is available at high frequencies. Many of them have markets with delivery dates in the future, as well as call and put options. Compared to markets like manufacturing, the market for commodities is much more developed.

Because they are mostly natural resources, the production process also tends to be unique. The extension of this paper will be focusing on farming products to explore the implications of a key feature of agricultural markets: the seasonality of their production process. By seasonality, I will include any production process that presents exogenous monthly variations throughout the year systematically every year. That is, seasonal production processes are variations that repeat themselves every year and hence can be forecastable within a certain range. Differences in seasonality allow for different future curves depending on the product in question, its harvest season and geographic location of markets.

Agricultural goods are different from most other goods since they can only be produced (harvested) during a certain interval of the year. Production decisions must follow a certain timing and schedule. This process is all common knowledge and allows for a set of future contracts to be signed before delivery dates. Since production is irregular but demand for these goods is steady throughout the year the inevitable answer is to add storage to the industry. Within a closed economy framework, the existence of storage means that throughout the year consumers will be "eating" old harvest leftovers (stored in silos) while waiting for the "new" harvest to come in at some point. Therefore future prices with delivery dates prior and after the new harvest will have different sources of supply. Since supply during different moments in time will come from different sources (harvests), then supply shocks (or information shocks regarding supply) should not "move" future prices in the same direction. Moreover, shocks regarding the new harvest (i.e., a USDA report that forecasts the next harvest) should only affect futures with post-harvest delivery dates whenever these are in backwardation, which is the usual case for agricultural commodities since it is common to see a "drop" in future prices that mature after the next harvest. However, data shows the opposite. Correlations between new and old future prices for all main agricultural commodities are not only positive but also consistently high throughout the years, usually between 0.7 and 1. The first explanation for this is that demand shocks might be autocorrelated, hence allowing for positive correlation between all future prices. However, Pirrong (2011) addresses this hypothesis as "incomplete" by analyzing the correlations under supply-side news shocks. Indeed, he finds that even after conditioning the sample to supply-related information shocks, correlation is still positive and close to one. Therefore, some "supply" based explanation is still required.

The main hypothesis in this paper is that new harvests are sold in both “old” and “new” markets. That is, if we start by assuming that harvests come “in pieces” (instead of all in one moment, as the literature does), market equilibrium will result in selling the “early part” of the harvest in the old markets, and the rest in the new one. Hence, both markets would have a common source of supply, allowing for supply induced correlation.

It is also important to say that the model used in this paper is consistent with standard literature. Competitive storage models have been widely used within the commodity literature, with many positive results when contrasted empirically. The main goal here is to explain correlations between future prices of seasonal goods without contradicting the key findings that competitive storage models have already achieved.

This paper is organized as follows: the present section is the introduction to the topic and main puzzle that this papers intends to address. Section 2 details the main findings of the literature and how this issue has been approached so far. Section 3 introduces the model, explains the puzzle and shows theoretical results. Section 4 presents empirical evidence for the assumptions used and results obtained. Section 5 concludes. An appendix is attached with proofs for all the results shown in section 3.

7 Literature overview

Storage theory has been widely used throughout the commodity literature. It dates back to almost a century ago with Kaldor’s 1939 convenience yield hypothesis stating that future prices are expected spot prices adjusted for storage costs, opportunity costs and an implicit benefit -convenience yield-. This theory contrasted with Keynes’s normal backwardation theory, which describes future prices as expected spot prices plus a risk premium. The empirical analysis of competitive storage theory however didn't come until the late 80's, probably due to the availability of data and computational capacity. Many statistical aspects of these markets were studied within this framework. Fama & French (1987) test storage models and analyze the relation between the basis (the difference between future and spot prices) and risk premiums (normal backwardation approach). They find evidence in support for storage theory and also, for some commodities, evidence in favor of the risk premium approach. In the same direction, Gorton, Hayashi and Rouwenhorst (2012) also find evidence for the predicted relation both between inventories and basis and between inventories and risk premiums in a much broader set of commodity data. They don't however find a relation between trading position and risk premiums. Both Ng-Pirrong (1994) and Fama & French (1988) analyze variability in spot and future prices and their relation with inventory levels, comparing model predictions with data, and again finding interesting results in its favor.

In a seminal paper, Deaton-Laroque (1992) apply standard rational expectations competitive storage model to thirteen commodities and match moments in the data like skewness, kurtosis and conditional variances depending on the type of demand shocks simulated. Dvir-Rogoff (2014) apply the storage model with growth to the oil market, obtaining results that match data in certain aspects for both before and after 1973, when supply became restricted.

More related to the issue addressed in this paper, Pirrong examines correlations between new and old harvest prices for six agricultural commodities for a thirty-year sample and shows a systematical high and positive correlation for all goods. Moreover, he repeats this analysis but then conditioning the sample on USDA report harvest forecast release dates, that is, he only tests for correlations on days where such reports are released. These USDA harvest reports contain information regarding supply: harvests forecast, quality conditions, expected timing, etc. These results conditioned on report release dates still show a high positive correlation. Therefore, one cannot simply explain such correlation through autocorrelated demand shifts as in Deaton-Laroque (1992). He demonstrates that these results are inconsistent with simulated results from a storage model and also claims that alternative explanations such as intertemporal substitution and inventories as inputs cannot obtain the quantitative desired results. Finally, he suggests that multiple commodities in a general equilibrium might be a more fruitful approach, but faces computational and quantitative constraints. The focus of this paper will be to reconcile this discrepancy between model and data.

8 Model

Time is continuous and runs from zero to some finite time T, so the economy starts at period 0 and ends at T. Initially, there is an endowment of stored goods $Q_0^s = Q(0)$, also known as "carry-in". Later, during the time interval $[a, b]$, $\quad 0 < a < b < T$ a harvest of y quantities (exogenous) comes into the market following some frequency $g(t)$ with domain $[a, b]$. Harvest scale y is random, hence it will be the only source of uncertainty. Frequency $g(t)$ is deterministic and perfectly known.

There will be two agents: consumers and speculators. Consumers buy good c for final consumption. Speculators on the other hand are risk neutral agents that wish to maximize profit. They buy goods in order to store and sell them later.

8.1 Consumer's problem

There is a representative agent that can consume either commodity good C or numeraire good w. Since the commodity good is a very small fraction of the economy, I will assume preferences are quasilinear in such good. Formally, her problem is:

$$\max E_0 \left[\int_0^T \exp(-rt) \left(Z\frac{C(t)^{1-\rho}}{1-\rho} + w(t) \right) dt \right]$$

s.t. $\quad S(t)C(t) + w(t) = \bar{W} \quad$ where C indicates consumption of the commodity good, w is the numeraire and $\bar{W}$ is income

FOC for consumers imply:

$$S(t) = \exp(-rt) Z C(t)^{-\rho} \tag{1}$$

8.2 Speculator's problem

There is a representative speculator, which stands for both farmers and financial agents in practice. Speculation occurs by buying goods for storage and selling them later. At each time t the speculator solves the following problem:

$$\max_{X(t,i)} E_t \left[\int_t^T m_t(i) S(i) X(t,i) di - S(t) X(t) \right]$$

$X(t,i) \geq 0$, for all t, i

$X(t) = \int_t^T X(t,i)di$

where $S(.)$ is the spot price, $X(t,i)$ is the good stored at time t to be sold at time $i > t$, $X(t)$ is total goods stored at time t and $m_t(i)$ is the stochastic discount factor between time i and t.

Optimality conditions imply that:

$$E_t\left[m_t(i)S(i)\right] \leq S(t) \quad , \text{ with } " = " \text{ if } X(t,i) > 0 \tag{2}$$

That is, in equilibrium the expected discounted value of the good has to be lower or equal than the spot price. The reason of the inequality is given by the non-negative constraint on storage. If expected discounted prices are greater than present ones, then arbitrage profits can be made, but the opposite is not the case, since you cannot have negative goods. Notice that if equation 2 holds with equality then speculators are indifferent between any amount of storage. In equilibrium stored amounts will be given by market clearing conditions:

$$X(t) = Q(t) + y(t) - D(S(t)) \tag{3}$$

Also, since $\dot{Q}(t) = y(t) - D(S(t))$, then

$$\dot{Q}(t) = X(t) - Q(t) \tag{4}$$

Also, since $\dot{Q}(t) = y(t) - D(S(t))$, we have

$$\dot{Q}(t) = X(t) - Q(t) \tag{5}$$

Lemma 1 *If storage is positive, consumption will move according to the following equation:*

$$\frac{\dot{c}(t)}{c(t)} = -\frac{r}{\rho} \tag{6}$$

The previous lemma states that consumption should move smoothly according to storage costs and elasticity. This will imply that prices move in a way such that they match interest rates. Hence, the solution be will be composed of intervals during which non-negative constraints are active or non active. It is trivial to point out that when the non-negative constraint is active, then consumption is just equal to harvest pick-up $c(t) = y(t)$. During intervals of with positive storage, consumption path is set by equation 6.

Take the case of some interval $[\alpha, \beta]$ with $Q(t) > 0$.The level of consumption can be found by solving 6:

$$C(t) = \exp(-\tfrac{r}{\rho}t)C(\alpha) \qquad \text{for } \alpha \leq t$$

where $C(\alpha)$ indicates some initial level of consumption determined by the feasibility condition:

$$Q(\alpha) + \int_{\alpha}^{\beta} y(i)di = \int_{\alpha}^{\beta} C(i)di$$

$$Q(\alpha) + \int_{\alpha}^{\beta} y(i)di = \int_{\alpha}^{\beta} \exp(-\tfrac{r}{\rho}t)C(\alpha)di$$

$$Q(\alpha) + \int_{\alpha}^{\beta} y(i)di = C(\alpha)\tfrac{\rho}{r}\left[1 - \exp(-\tfrac{r}{\rho}(\beta - a))\right]$$

$$C(\alpha) = \frac{\frac{r}{\rho}}{\left[1 - \exp(-\frac{r}{\rho}(\beta - \alpha))\right]}\left[Q(\alpha) + \int_{\alpha}^{\beta} y(i)di\right] \quad \text{for some interval } [\alpha, \beta] \tag{7}$$

8.3 Futures market

Additional to storage, speculators can sign in any moment future contracts that set prices in a specific time. By non-arbitage conditions, these should have the following value:

$$\exp(-r(t + j))F(t, t + j) = E(m_t(t + j)S(t + j)) \tag{8.a}$$

where $F(t, t+j)$ indicates a future contract signed at time t, regarding the commodity price at moment $t+j$. We can rewrite equation 6 under some equivalent measure to obtain:

$$\exp(-r(t+j))F(t, t+j) = \hat{E}(S(t+j)) \tag{8.b}$$

Equations 8 state that the value of future contracts signed in t depends on expectations over spot prices at delivery moment $t+j$. If this condition was not the case, then speculators would go long/short and make infinite expected profits.

8.4 Centralized problem

The previous solution can also be mapped into a planner's problem, hence proving that it is also efficient. In summary, the planner's problem is:

$$\underset{x}{Max}\ E_0 \left[\int_0^T \exp(-rt) Z \tfrac{C^{1-\rho}}{1-\rho} dt \right]$$

st $\quad \dot{Q} = y - x \quad Q \geq 0 \quad Q_0^s > 0 \quad$ where $y(t)$ indicates the harvest that is picked up in moment t.

$Q(T) \geq 0$

The solution will depend on harvest variable $y(t) = G(t)H$, where $G(t) = \int_0^t g(i)di$ indicates the fraction of harvest H picked up up to time t. $Q(T)$ indicates storage leftover at time T. Optimality implies that $Q(T) = 0$, since if not $(Q(T) > 0)$ we could find a better solution by just reallocating such storage to consumption.

As can be seen in appendix, the solution will have for every t :

$$\lambda = \exp(-rt) Z c^{-\rho} \tag{9.a}$$

$$\lambda' = -H_Q = \lambda \frac{r}{\rho} + \eta \tag{9.b}$$

$$\dot{Q} = y - x \tag{9.c}$$

$$\eta Q = 0 \tag{9.d}$$

where λ is the multiplier from the Hamiltonian related to the law of motion (also interpreted as "shadow prices") and η is the multiplier associated with the non-negative constraint.

Therefore, for any social planner's solution I can find a decentralized version by setting $S(t) = \lambda(t)$ to find prices and use non-arbitrage conditions to construct the future curve.

8.5 Degenerate harvest case

Let's assume now that in moment J, harvest y (random) will come in all at once. Mathematically, we assume that $g(t)$ degenerates in a single point $0 < J < T$, that is, $g(J) = 1$ and $" = 0"$ otherwise.

An easy way to frame the solution to this situation is as if there were two markets, one that goes from $[0, J)$and has carry-in Q_0^s and another one during interval $[J, T]$.with carry-in y. Obviously, arbitrage between these two markets in time is possible with storage technology (i.e. bring present goods to the future). An intuitive approach is to call the first term "old-harvest term" and the second one "new harvest term".

Hence there are two possible scenarios, depending on the size of y. For simplicity I will assume y can be in two ergodic states: Y_H , where y is always big enough to cause a drop in future prices (see scenario II below), and Y_L, where y is never big enough to cause a drop in future curves (scenario I). The goal of this is assumption to simplify the analysis even though it doesn't necessarily correspond to reality. By focusing on the simple case, I can concentrate in cases where backwardation takes place, and yet still get the desired correlation found in the data. Hence results will be strong since

I will not need to go for possible "jumps" for contango to backwardation as a main explanation for correlation.

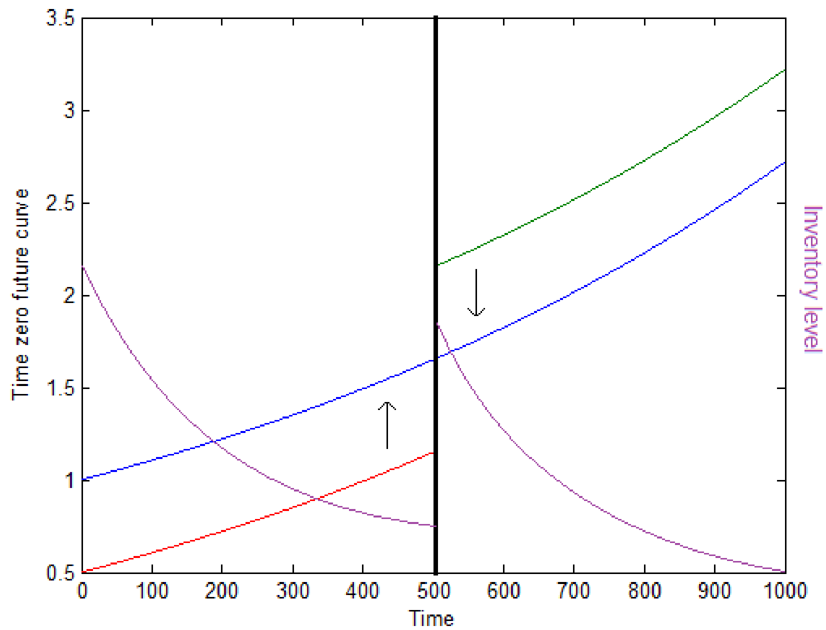
Scenario I: $\hat{E}\left[S(J^-)\right] = \hat{E}\left[S(J^+)\right]$. One interval all along $[0, T]$ with positive storage

Proposition 1 *If the harvest is not big enough, there will be no "kink" in prices, since there are incentives to have positive storage all along the whole time interval* $[0, T]$

Proof. See appendix ■

Here it is common knowledge at time zero that $y \in Y_L$, hence, y will always be small.

Figure I



Here, speculators have incentives to arbitrage by reducing available goods in the first term and storing them for the second one, hence increasing old-harvest prices and reducing new-harvest ones. This will continue up to the point where they converge, only then will we be in equilibrium (blue line). Inventories will drop throughout time but never reach zero until the end of the period. For this to happen, the harvest must be small enough. I call this the "small harvest effect".

Now we will have a single interval solution, despite the harvest arriving at J. Now both carry-in Q_0^s and harvest y will be consumed smoothly throughout $[0, T]$. Since y is ex ante a random variable, in equilibrium the price will be:

$$S_0 = M(0,T)\hat{E}_0\left[(Q_0^s + y)^{-\rho}\right]$$

$$F(0,t) = M(t,T)\hat{E}_0\left[(Q_0^s + y)^{-\rho}\right] \qquad 0 \le t \le T$$

where $M(A,B) = \exp(-r(A))Z\left[\frac{\frac{r}{\rho}}{\left[1-\exp(-\frac{r}{\rho}B\right]}\right]^{-\rho}$

The results are pretty intuitive. The first term $M(t,T)$ is a constant that depends on interval length "$T-0$" and delivery date t and is decreasing in T. That is, the bigger the interval, the higher the price, since you have to stretch storage throughout more time. Notice also that prices are continuous and decreasing in carry-in variable Q_0^s (the more reserves you start with, the more you can consume in each moment of time).

Notice that in scenario I, changes in expectations of y change prices all over the interval, even before the harvest. Here the correlation among futures with both old harvest and new harvest delivery dates should be positive and close to 1. This matches what happens in the data. The puzzle however, comes when prices are in backwardation, as will be seen in scenario II. Indeed, the data shows that correlations are still high despite future curves show very often a "backwardation" position.

From this point onwards I will define as a "Hotelling interval" any time interval where:

a) Future curves move according to interest rates (plus implicit storage costs, in this case zero)

b) Storage is positive

c) No goods are "left over" once this interval is over

Hence, scenario I shows a future curve that presents a "Hotelling" interval all along the entire time framework $[0,T]$.

Scenario II: $\hat{E}\left[S(J^-)\right] > \hat{E}\left[S(J^+)\right]$. Two intervals $[0,J)$ and $[J,T]$, with $Q(J) = 0$.

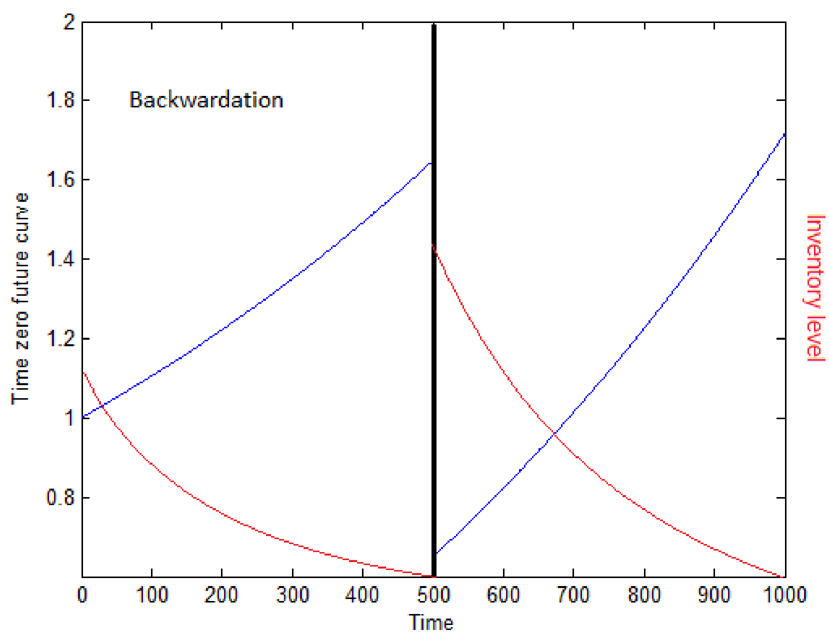
Proposition 2 *If the harvest is big enough, there will be a negative drop in prices at time J, since there are no incentives to store for after the harvest.*

Proof. See appendix ■

The intuition is the following. If agents know that a big harvest is to come at time J, then it is optimal for them to consume all their stored goods right to the moment before the harvest comes in. Then, for $t \geq J$, they can consume whatever the resulting harvest was.

In this case, I assume it is known at time zero that $y \in Y_H$, hence, y will always be big enough such that we are in backwardation. Therefore, the future curve will take the shape of figure II:

Figure II



Here, by non-negativity of storage, arbitrage is not possible and we have an equilibrium with a negative discontinuous jump in J. Inventories are fully used before the harvest arrives at J. This will happen if the harvest is big enough. I call this the "big harvest effect".

Mathematically, at time 0, future prices will be

For the first term $\quad 0 < t < J,$

$$S_0 = M(0, J) Q_0^{s^{-\rho}}$$

$$F(0,t) = M(t,J)Q_0^{s^{-\rho}}$$

For the second term $J \leq t \leq T$, since we have uncertainty over harvest size y, expectations will now be a part of the solution:

$F(J,J) = F(J,t)\exp(-r(t-J)$

$$F(J,J) = M(0,T-J)\hat{E}_0\left[y^{-\rho}\right]$$

$$F(J,t) = M(t-J,T-J)\hat{E}_0\left[y^{-\rho}\right] \qquad J \leq t \leq T$$

Figure II shows a future curve that breaks in J. Since storage cannot be negative, new harvest markets (futures that mature after J) will not be linked with old harvest ones (futures that mature before J). Hence if expectations regarding the harvest change, only the second term price curve should change. The first one should not move at all. Therefore, at least for shocks that are small, curves should not move together, resulting in close to zero correlation.

The puzzle here is that this is not what is seen in the data. Indeed, Pirrong shows that correlations between new and old harvest prices are positive and very high, for all main agricultural crops, even when most of the time future curves are in backwardation. An initial explanation that follows from Deaton-Laroque (1994) is that such correlation is explained by demand shocks that tend to have persistence. However, although it may be true, high correlations still hold even when controlling for information shocks regarding only supply variables. Pirrong (2015) tests the relationship between old harvest prices versus and new harvest ones by controlling for days when USDA reports (that contain information only related to supply) are released. He finds that the previous results still hold. Hence, there must be some supply side explanation for this feature of the data (not necessarily contradicting the persistent demand hypothesis). The intention here is to obtain such correlation within a competitive storage benchmark by adding a more realistic assumption: the fact that harvests appear continuously rather than in a one-time lump fashion.

8.6 Continuous harvesting

I will now proceed to the version of the model that drives the main results in this paper. The goal here is explain supply-side induced correlation even under backwardation conditions.

Let there be a period of harvesting in the interval $[a,b], 0 < a < b < T$. Within that period, harvest comes in according to a known frequency function $g(t)$, with domain $a \leq t \leq b$. The size of the harvest, H is a random variable, and information regarding such variable is revealed at some moment J, $t < J \leq a$.

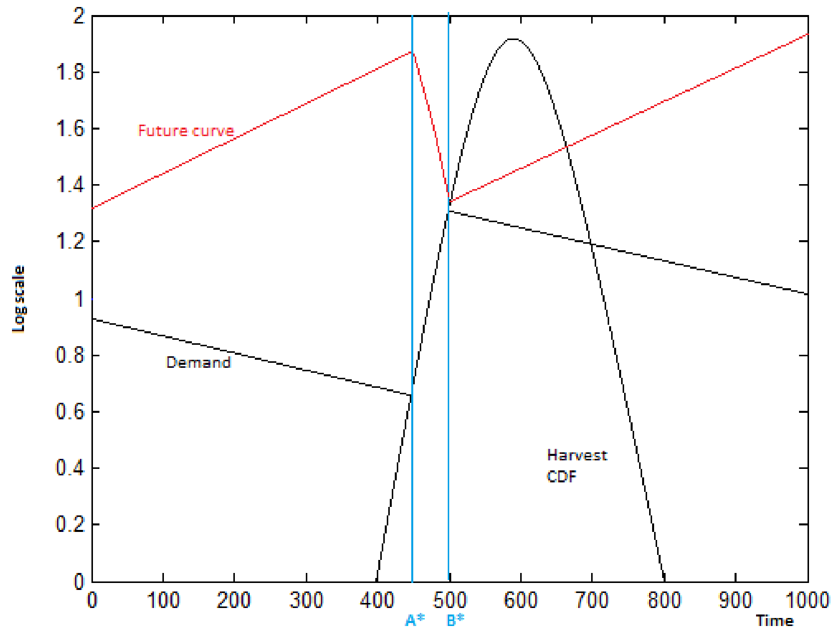
Therefore, at a given moment t, total goods harvested up to that point are $y(t) = G(t)H$, where $G(t) = \int_0^t g(i)di$. For simplicity, I will assume $g(a) = g(b) = 0$ ("smooth" starting and ending). Since the goal of the paper is to explain correlation under backwardation cases, I will assume that the distribution of H is such that probability of contango is zero. Hence we will be under backwardation almost surely. Since it is already known that under contango correlation is high and close to 1, this assumption will make results stronger and do not contradict in any sense other possible scenarios. Formally:

$H \in \left[\underline{\mathrm{H}}, \bar{H}\right]$, where $0<\underline{\mathrm{H}}<\bar{H}$, and $\underline{\mathrm{H}}$ is such that always for some $t \in [0,T]$, $\frac{\dot{S}(t|\underline{\mathrm{H}})}{S(t|\underline{\mathrm{H}})} < r$ for all t, $0 \leq t \leq T$.

Lemma 2 *If $\underline{H}$ is such that spot prices drop at some time, then this is also true for any $\underline{H}<H$.*

Proof. See appendix ∎

Figure III



Theoretical equilibrium future curves under no uncertainty

Claim 1: There exists an initial Hotelling interval $[0, A]$, with $A \geq a$.

Intuitively, if the early part of the harvest $G(A)$ is small, we will have the "small harvest effect" until some moment $A \geq a$. At that point, the major part of the harvest starts flowing in, causing a "big harvest effect", hence followed by a negative slope. If $g(a) = 0$ (i.e., harvest starts "smoothly"), then $A > a$.

Claim 2: At some moment B, $a \leq B < b$, storage will occur and hence Hotelling's rule will apply from that moment onwards until the ending of the period, T.

The logic operating in this case is that after b, we must have some stored goods, otherwise nothing would available and prices would skyrocket to infinity. Hence the storage process must start at some moment prior to the harvest ending, $B < b$.

Proof. See appendix. ■

Let us take the case of the figure III, where $A > a$. In this case, prices can be defined within three intervals : $[0, A), [A, B)$ and $[B, T]$

8.7 Main Result

During the first and the last intervals, Hotelling's rule applies, hence time zero future prices can be written as:

$$\begin{aligned} F(0,t) &= M(t,J)\left[Q(0)-\bar{Q}(J)\right]^{-\rho} \\ 0 &< t<J \end{aligned} \tag{10.a.1}$$

$$\begin{aligned} F(0,t) &= M(t,A-J)E_0\left[\left(G(A^*)H+\bar{Q}(J)\right)^{-\rho}\right] \\ J &\leq t<A \end{aligned} \tag{10.a.2}$$

$$\begin{aligned} F(0,t) &= M(t-B,T-B)\hat{E}_0\left\{[(1-G(B))H]^{-\rho}\right\} \\ B &\leq t\leq T \end{aligned} \tag{10.b}$$

Where $\bar{Q}(J)$ will be such that in equilibrium $F(0,J^-)=F(0,J^+)$. Hence $\bar{Q}(J)$ is a function of expectations and carry-in $Q(0)$.

During the interval $[A,B)$ Hotelling's rule is not active, hence there is no storage taking place. Therefore, demand is constantly supplied by the continuous harvest inflow. With continuous harvesting, the market clearing equation becomes:

$$ZS_t^{-\rho}=D(S_t)=dG(t)H \qquad \text{for } A\leq t<B$$

$$\begin{aligned} F(0,t) &= \hat{E}_0\left(S_t\right)=Z\hat{E}_0\left[(dG(t)H)^{-\rho}\right] \\ A &\leq t<B \end{aligned} \tag{10.c}$$

Intuitively, we can call equations 10.a.1, 10.a.2 "old harvest prices", 10.b and 10.c "new harvest prices", as the literature does.

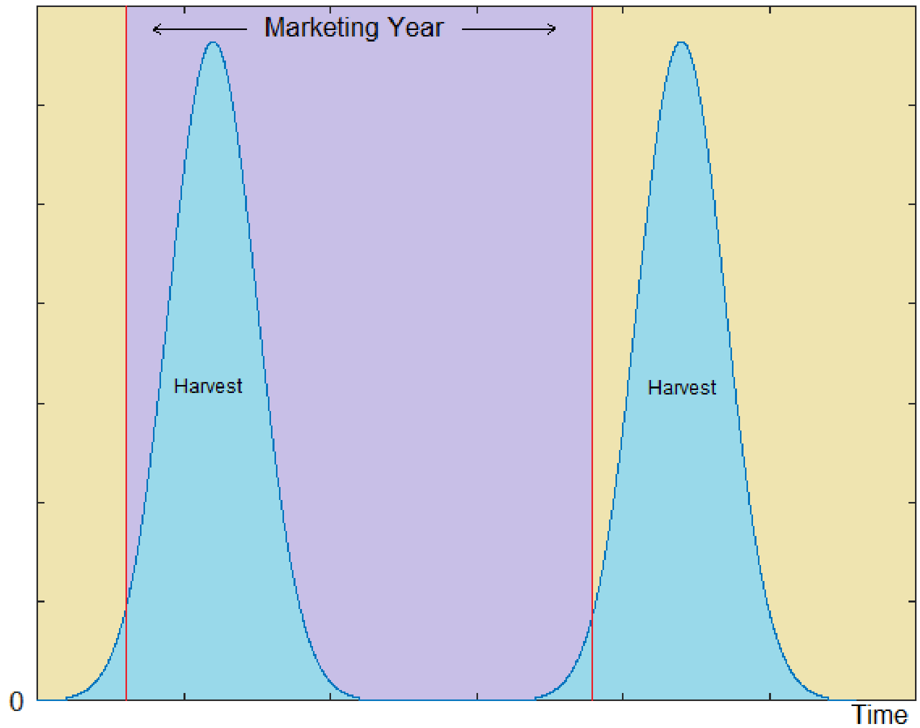
The only source of randomness comes from $E_0\left[\left(G(A^*)H+\bar{Q}(J)\right)^{-\rho}\right]$ in 10.a.2, $\bar{Q}(J)$ (that depends itself on expected harvest – see appendix) in 10.a.1, $\hat{E}_0\left(H^{-\rho}\right)$in 10.b and 10.c. Clearly, all terms have the variable H in common and are strictly decreasing in it for any $\rho>0$. Prices are set according to expectations over those terms. If expectations change (due to new information for example) then both new and old harvest prices will change in a similar way. Correlation will be high. Hence, the model solves the puzzle presented by Pirrong without abandoning the inventory based framework.

The source of randomness is the set of possible conditional expectations agents can have at each time t. Hence, the final random variable is the conditional expectation itself. Depending on different sets of information, expectations in time zero (or time $j>0$, WLOG) will be different.

Finally, if we take a closer look at equation 10.A, we can see that the expectations term has both carry-in $Q(0)$ and early harvest $G(A)H$. Hence, I proceeded to simulate correlations for different relative values of carry-in to expected early harvest, finding that as long as the early part of the harvest $G(A)H$ is positive, correlation will always be between 0.7 and 1, as in the data. Results are robust to different values of elasticity ρ. Hence, the model successfully achieves the values observed in the data, even after conditioning the possible set of results to backwardation future curves only.

Results can be summarized in Figure IV. The concept is that, given that harvest gets picked up continuously, "Marketing years" for seasonal commodities are defined endogenously by containing both old harvests and "early-parts" of new ones. Therefore, in each marketing year, there are always goods supplied by the new "upcoming harvest, allowing for positive strong correlation between future prices with different delivery dates.

Figure IV



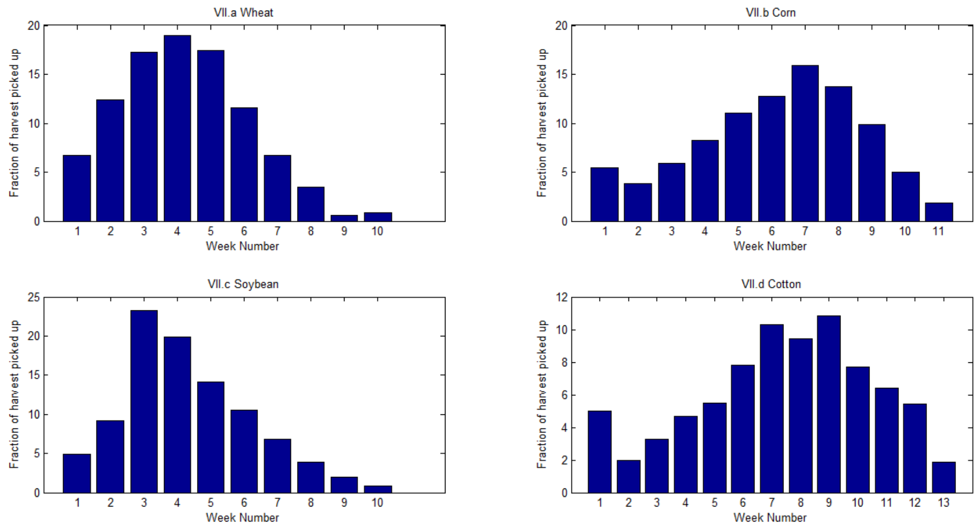
Marketing Year defined endogenously by markets

9 Empirical evidence

9.1 Hypothesis

A key component of the model used in this paper is the frequency and interval through which the harvest is "picked up". Hence, it is of high relevance that such assumptions match the data. Figures V (a-b-c-d) summarize the average harvest pick-up rate for 2011-2016 in terms of weeks.

Figures V



Source: USDA. Average 2011-2016 fraction picked during each harvest week.

Two main points must be driven from figures V. First, harvests for all main commodities follow some bell shaped distribution throughout time. Second, the harvest interval lasts between 10 to 13 weeks, i.e. between one fifth and one fourth of a calendar year. Therefore, not only does harvest progress come in a bell shaped form, but it also requires a significant amount of time. That is, the size of the interval is relevant.

9.2 Price and Inventory Seasonality

The model predicts that prices and inventories should both follow a heavy seasonal pattern. Moreover, between harvests, we should see prices grow as a function of interest rates plus storage costs (Hotelling interval). Figures VI.a, b c and d depict seasonality components (monthly dummies) for first generic futures for corn, soybean, wheat and cotton. Also, as a comparison, I show future curves for non-seasonal commodities like gold and silver. First generic futures reflect the value of the future contract with closest delivery date. Once that date is reached, the series changes to the following one, and so on. Hence, they are not spot prices, but rather an approximation that might be two to three months lagged. An example of future curves can be seen in Figures VII instead. In these, Hotelling intervals are pretty clear, and the relation between harvest season and backwardation points is also evident.

Figure VI.a

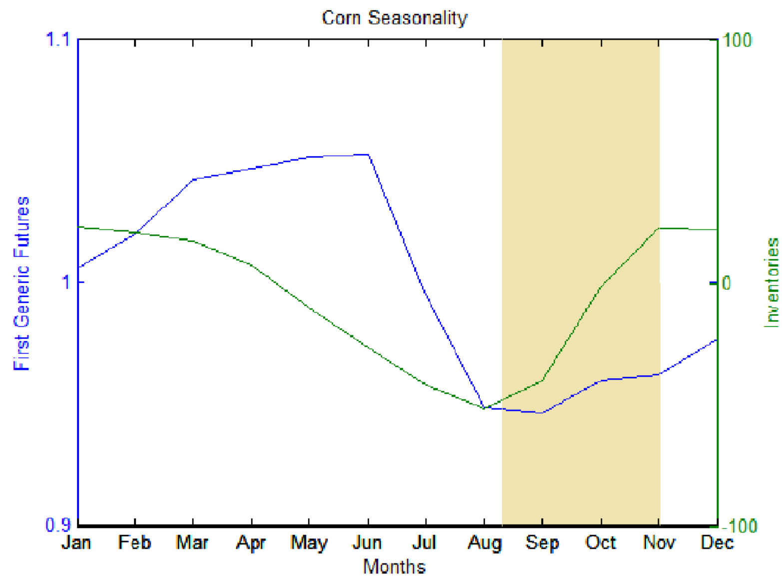


Figure VI.b

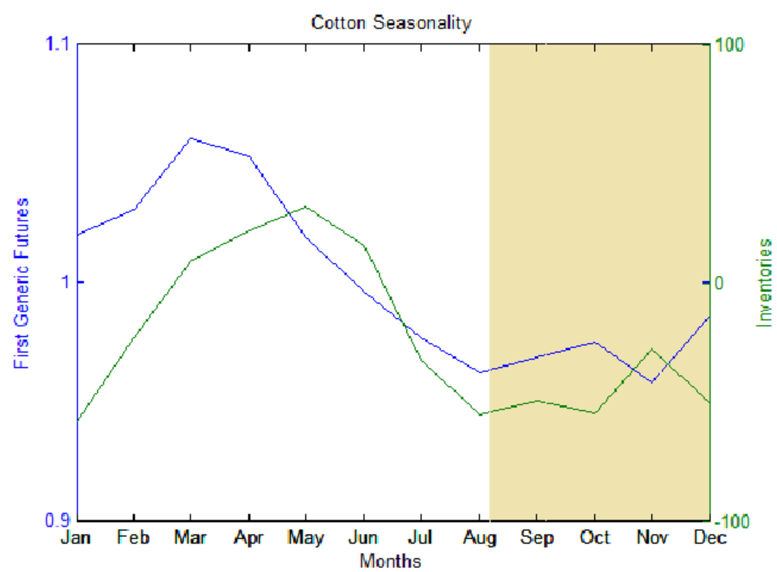


Figure VI.c

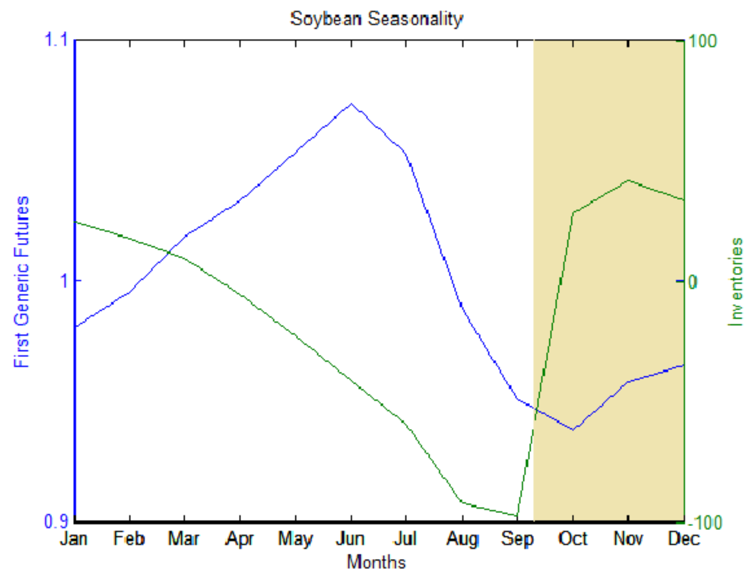
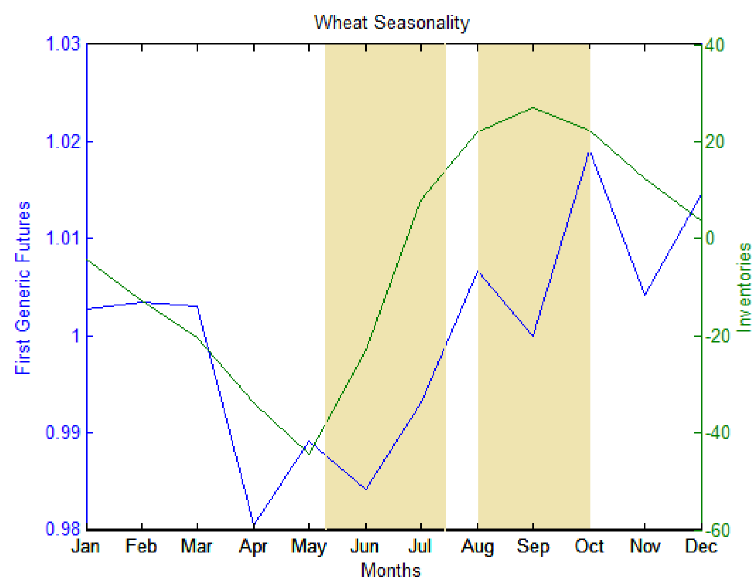


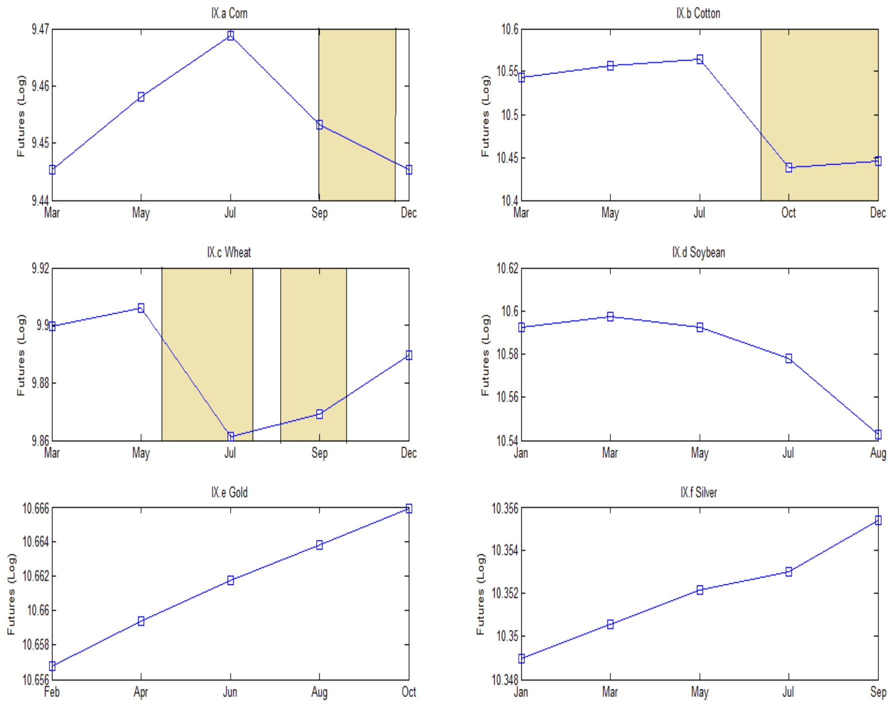
Figure VI.d



Source: Bloomberg, USDA and Gorton & Hayashi (2012). Future's month seasonality is calculated for first generic futures with moving average multiplicative methods. Inventory levels are monthly log deviations from past 12 month averages. Shaded areas indicate harvest periods.

Figures VII

Future curves by Jan-02-2004



Source: Bloomberg and USDA. Horizontal axis indicate delivery dates

It is important to point out that when looking at future curves we only observe five months. We cannot know what equilibrium values are between such dates. The lines drawn in figures VII are merely illustrative, they do not indicate prices for moments between delivery dates (squares). Additionally, it is useful to point out that wheat has two harvests: spring (65% of total) and winter (other 35%).

9.3 Finding the revelation moment

In the model I assume that at some moment J information is revealed. Of course, in reality that is not necessarily true: there may be more than one or even a continuous of "J"s. The important thing is that when information is revealed consumption paths are corrected and hence equilibrium prices change. Therefore, we should expect to see higher volatility during these events. Table I shows some suggestive evidence. The dependent variable is the standard deviation of generic futures during the last 60 business days (rolled over throughout the years 2002 - 2017). Besides the agricultural commodities of interest, I add gold, silver and oil to have an outside reference. Indeed, it is clear that during July, August and September volatility increases significantly, coinciding with pre- harvest months for all crops. This pattern is not observed with the other three non-agricultural commodities.

Table I

	Dependent variable: Standard Deviation of Generic future - past rolling over 60 days						
	Corn	Soybean	Wheat	Cotton	Gold	Silver	Oil
Feb	0.00249	.002498	-0.0003201	.00009	-.0005783	-.00033	-.00757**
	[.0022]	[.0022]	[.0022]	[.0026]	[.0012]	[.0023]	[.003]
Mar	-.0034854	.006782***	-0.0031457	.003319	-.0014	.00459**	-.01558***
	[.0021]	[.0021]	[.0021]	[.0025]	[.0012]	[.0022]	[.0029]
Apr	-.00607***	.007844***	-0.0047109**	.00247	-.005377***	.00434*	-.008***
	[.0022]	[.0021]	[.00222]	[.0025]	[.0012]	[.0023]	[.0029]
May	-.00483**	.004992**	-0.004672**	.003377	-.003592***	.00967***	-.0092***
	[.0022]	[.0021]	[0.00224]	[.0025]	[.0012]	[.0023]	[.003]
Jun	-.0008705	.007947***	0.0008894	.00637**	-.003836***	.00575**	-.00888***
	[.0022]	[.0021]	[.0021]	[.0026]	[.0012]	[.0023]	[.003]
Jul	.02137***	.01511***	0.015532***	.01721***	-.00723***	-.00063	-0.01683***
	[.0022]	[.0021]	[.00222]	[.0025]	[.0012]	[.0023]	[.0029]
Aug	.0449***	.0283***	0.0225113***	.02175***	-.007425***	-.002736	-.01745***
	[.0021]	[.0021]	[.00222]	[.0025]	[.0012]	[.0022]	[.0029]
Sep	.02847***	.0332***	0.0123016***	.0063034**	-.003526***	.00601***	-.01506***
	[.0022]	[.0021]	[.00222]	[.0025]	[.0012]	[.0023]	[.003]
Oct	.01477***	.03109***	0.00734***	.0051884**	-.0003	.01378***	-.01458***
	[.0021]	[.0021]	[.00222]	[.0025]	[.0011]	[.0022]	[.0029]
Nov	.02149***	.0255***	0.004888**	.01386***	-.001	.01078***	-.0032
	[.0022]	[.0021]	[.00222]	[.0025]	[.0012]	[.0023]	[.0029]
Dec	.00759***	.0103***	0.0013325	.007011***	.00003	.0033	.0013
	[.0021]	[.0021]	[.00222]	[.0025]	[.0012]	[.0023]	[.0029]
cons	.04393***	.03409***	0.0525291***	.0472***	.0352***	.0527	.0729
		[.00151]		[.0018]	[.00083]	[.00161]	[.0020967]
R-sq	0.25	0.15	0.08	0.04	0.03	0.03	0.03
Obs	3,743	3,743	3,743	3,743	3,743	3,743	3,743

*p<0.1, **p<0.05, ***p<0.01

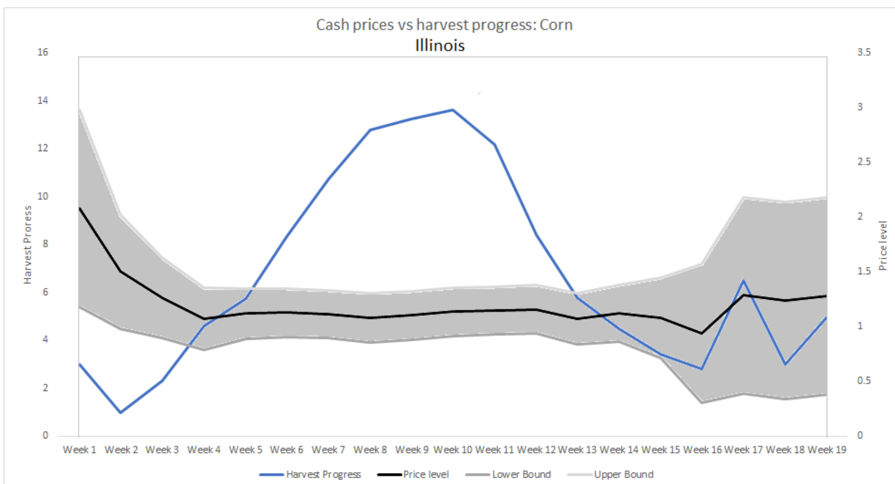
Source: Bloomberg. First generic futures. Jan 2002- April 2017

The case of wheat deserves a special explanation. April/May is planting season for winter wheat, that only accounts for 35% of total US wheat production. However, since spring wheat is the only major crop planted in October/November, and this coincides with corn, winter wheat and soybean harvest dates, it is logical that whatever economic decisions are made in April determine the amount to be planted later on. That is, once we know how much area has been dedicated to corn, soybean and winter wheat, the area to be used for spring wheat is straightforward.

9.4 High frequency data

Figure V depicts a theoretical future curve with continuous delivery dates. In reality, future markets show only certain points throughout the year, with very low frequency. Therefore, I proceed to work with cash prices, since they are available both daily and weekly In the model future values depend on expectations over delivery date spot prices. Therefore, I proxy such curves with (average) spot price series:

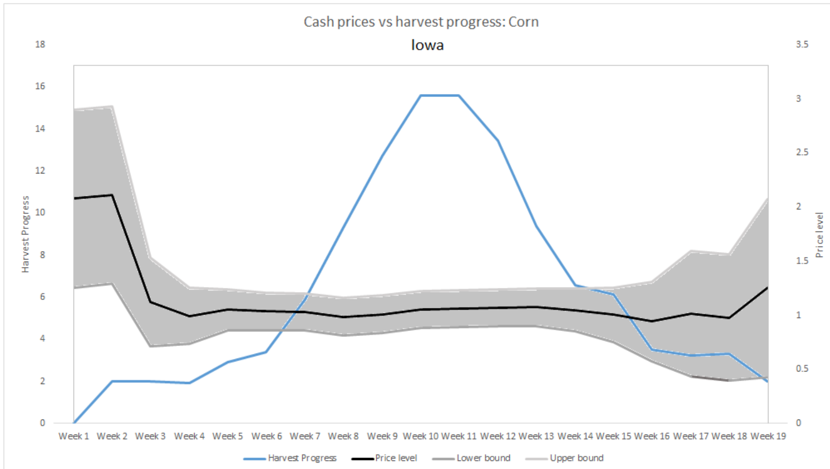
Figure VIII.a



Source. USDA. Lines depict 2000-2016 weekly average

Two standard deviation interval. Price level in log values

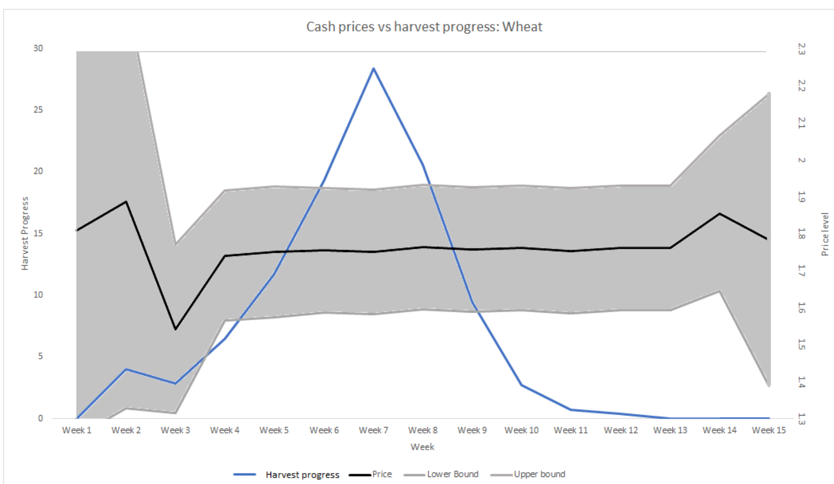
Figure VIII.b



Source. USDA. Lines depict 1992-2016 weekly average

Two standard deviation interval. Price level in log values

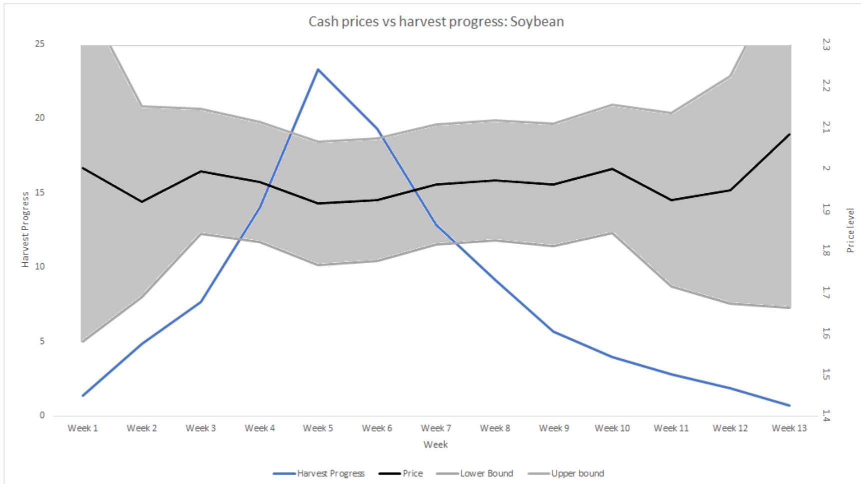
Figure VIII.c



Source. USDA. Lines depict 1992-2016 weekly average for Kansas

Two standard deviation interval. Price level in log values

Figure VIII.d



Source: USDA. Lines depict 1992-2016 weekly average for Illinois

Two standard deviation interval. Price level in log values

Figures VIII.a, VIII.b, VIII.c and VIII.d show average changes in prices and harvests for 1992-2016 throughout the harvest weeks, for corn, winter wheat and soybean in the US. That is, I take the average harvest pick-up of each week of the harvest calendar and compare it with the average level of prices (in logs).The following two points come out. First, during the first weeks prices already drop, even before the vast part of the harvests starts. Second, price patterns stabilize before the harvest peak is reached. Both facts match the models predictions and follow the theory of storage's prediction that agents anticipate to future markets and store for later, rather than selling everything today at perhaps a lower price.

10 Conclusions

Competitive storage theory is a widely used framework when it comes to commodity markets. Although it has been proven successful in many ways, there are still some issues that remain unexplained. Particularly, the standard theory's prediction regarding correlation between futures with pre and post harvest delivery dates does not match empirical evidence. In this paper I address this main issue without altering the fundamentals of storage models. by introducing continuous harvesting, rather than the "at once" endowment approach. This allows me to separate intervals of the same harvest between "early part" and "non early". Results are perhaps best perceived in Figure IV. Through non-arbitrage mechanisms, I manage to link in equilibrium old(pre) harvest prices with the early fraction of the upcoming harvest. Hence, the same source of supply (the upcoming harvest) is present in both old and new markets, allowing for the supply-side explained high correlation found in the data. Therefore, what the model is showing is that the marketing year is defined endogenously by markets by splitting harvests between early and non-early. Previous models could not capture this feature because harvests were usually defined as a "all-in-one-moment" drop. The innovation here is that giving length and continuity to harvest pick-up allow for this key aspect.

Results are robust to different parametrization of harvest size and demand elasticity, and are valid even when markets feel they will be in backwardation. As long as the "early harvest" (i.e., the fraction of the harvest sold in old markets) is positive, correlation will be positive and significantly high. These results are valid for any goods that are seasonally produced, therefore not constraining results to any particular agricultural commodity. I show empirical evidence to support the hypothesis of a continuous bell shaped harvest pick-up. I then take the case of corn, soybean, wheat and cotton, four commodities where the US leads in many aspects (consumption, production, exports), and contrast it with the theoretical predictions. I find very similar results to those given in the model, for both low and high frequency prices (futures and cash prices, respectively).

Assumptions used and simulations were designed to explicitly explain and show a supply-side source of correlation. Supply was assumed to be exogenous, demand to be deterministic and markets to take place within a closed economy framework. This by no means disregards other type of shocks that take place in reality, but merely intends to disentangle a remaining puzzle within the competitive storage literature, the fact that correlations between futures have not yet been explained from a supply point of view.

11 References

Aviv, Y., A. Pazgal. 2008. Optimal Pricing of Seasonal Products in the Presence of Forward-Looking Consumers. Manufacturing & Service Operations Management 10 (3) 339-359

Brennan M. (1958). "The Supply of Storage." The American Economic Review, 48,50-72.

Brennan, Donna C., Williams Jeffrey C., and Wright, Brian D. "Convenience Yield Without The Convenience: A Spatial-Temporal Interpretation of Storage Under Backwardation, " Economic Journal, 1997, 107(443), pp. 1009-23.

Colin A. Carter and Cesar L. Revoredo Giha."The Working Curve and Commodity Storage under Backwardation".American Journal of Agricultural Economics Vol. 89, No. 4 (Nov., 2007), pp. 864-872

Deaton, Angus, and Guy Laroque, 1992, On the behavior of commodity prices, Review of Economic Studies 59, 1–23.

Deaton, Angus, and Guy Laroque, 1996, Competitive storage and commodity price dynamics, Journal of Political Economy 104, 896–923

Dixit, A. and Pindyck, R. (1994) Investment under Uncertainty. Princeton University Press.

Dvir, Eyal, and Kenneth Rogoff. 2014. "Demand Effects and Speculation in Oil Markets: Theory and Evidence." Journal of International Money and Finance 42 (April): 113-128.

Fama, Eugene and Kenneth French (1987), "Commodity Futures Prices: Some Evidence on Forecast Power, Premiums, and the Theory of Storage,"Journal of Business 60(1), 55-73.

Fama, Eugene and Kenneth French (1989), "Business Conditions and Expected Returns on Stocks and Bonds," Journal of Financial Economics 25, 23-50.

Gorton, Gary, Fumio Hayashi and K. Geert Rouwenhorst (2013), "The Fundamentals of Commodity Futures Returns," Review of Finance 17 (January 2013), 35-105.

Gustafson, Robert L. Carry Over Levels for Grains: A method for Determining Amounts that are Optimal Under Specified Conditions. Technical Bulletin 1778, Washington D.C.: U.S. Department of Agriculture, 1958.

Kaldor, N. (1939). "Speculation and Economic Stability." Review of Economic Studies, 7, 1-27.

Keynes, John M., 1930, A Treatise on Money (Harcourt Brace, New York)

Muth, John "Rational Expectations and the Theory of Price Movements." Econometrica, 1961, 29(3), 315-35.

Ng, Victor K., and S. Craig Pirrong, 1994, Fundamentals and volatility: Storage, spreads and the dynamics of metals prices, Journal of Business 67, 203–230.

Pindyck, Robert (1994), "Inventories and the Short-Run Dynamics of Commodity Prices", RAND Journal of Economics 25 (1): 141-159.

Pindyck, Robert. (2001). The Dynamics of Commodity Spot and Futures Markets:A Primer." Energy Journal 22, 3: 1-30.

Pirrong, S. Craig, 1998, Price dynamics and derivative prices for continuously produced, storable commodities, Working paper, Washington University.

Pirrong, S. Craig, 2011. Commodity Price Dynamics: A Structural Approach. Cambridge University Press.

Samuelson, P. "Intertemporal Price Equilibrium: A prologue to the Theory of Speculation, "Weltwirtschaftliches, Archiv, 1957, 79(2), pp.181-219.

Samuelson, P. (1971): "Stochastic Speculative Price," Proceedings of the National Academy of Sciences, 68, 335-337.

Schwartz, Eduarto S. (1997), "The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging", Journal of Finance 52: 923-973.

Telser, Lester G. "Futures Trading and the Storage of Cotton and Wheat," Journal of Political

Economy, 1958, 66(3), pp. 233-255.

Turnovsky, Stephen J. “The Determination of Spot and Futures Prices with Storable Commodities,” Econometrica, 1983, 51(5), pp. 1363-87.

Working, H (1949). “The Theory of the Price of Storage.” American Economic Review, 39, 1254-1262.

Williams, J. (1986). “The Economic Function of Futures Markets.” Cambridge: Cambridge University Press.

Williams , J. and B. Wright (1989). “A Theory of Negative Prices for Storage.”Journal of Futures Markets, 9, 1-13.

Williams , J. and B. Wright (1991). “Storage and Commodity Markets.” Cambridge: Cambridge University Press.

Part III

Drivers of Inflation in LATAM

Francisco Arroyo Marioli, UCLA

Vibha Nanda, IMF

Frederik Toscani, IMF

In this section we analyze empirically inflationary processes in major Latin American countries: Peru, Chile, Colombia, Mexico and Brazil. We proceed to decompose core inflation time series into several drivers by estimating augmented versions of Phillips curves using data from the early 2000s until 2018. We find that domestic factors, such as persistence and expectations are still the main drivers. Foreign factors such as trade partner inflation and exchange rates can play an important role too. Finally, GDP cycles are still significant but rather small in explaining overall inflation processes.

12 Introduction

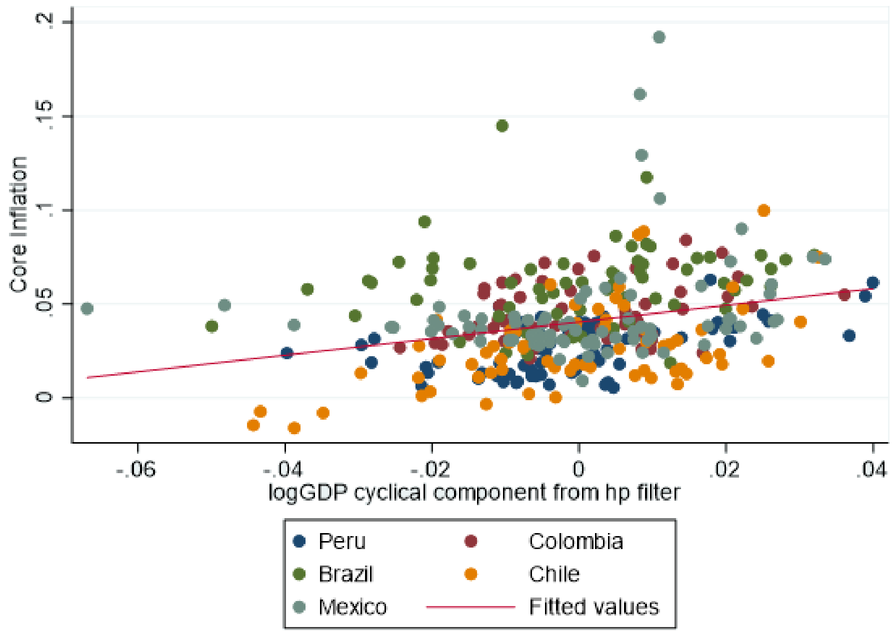
The last 20 years has been a period of inflation stabilization within most Latin American (LATAM) economies. Particularly, most countries have followed inflation targeting schemes, with significant success. This has brought several concerns on wether Phillips curves (PCs) may have "flattened", in the sense that inflation is no longer sensitive to output gaps. Moreover, given an increasing role of international trade it could be the case that external factors play now a bigger role in local inflationary processes. Many authors have explored this issue for many economic areas and countries,.such as the US (Abdih et al, 2016), the Euro Area (Abdih et al, 2018), Colombia (Lanau et al, 2018) and Chile (Naudon and Vial, 2016). Though many of the results are on a case-by-case basis, some common findings emerge, such as a reducing slope for Phillips curves and an increasing role for foreign factors in inflationary processes. In this paper, we address the same issue and focus on Latin America's biggest five economies -excluding Argentina. More formally, we intend to revisit Phillips curves for LATAM by including foreign factors such as exchange rates and trading partner's inflation in our analysis.

In summary, we find that domestic factors still play a major role in determining inflation. Past inflation and expectations still seem to be the main drivers behind inflationary processes, with economic slack playing still a role but smaller.

12.1 Stylized Facts

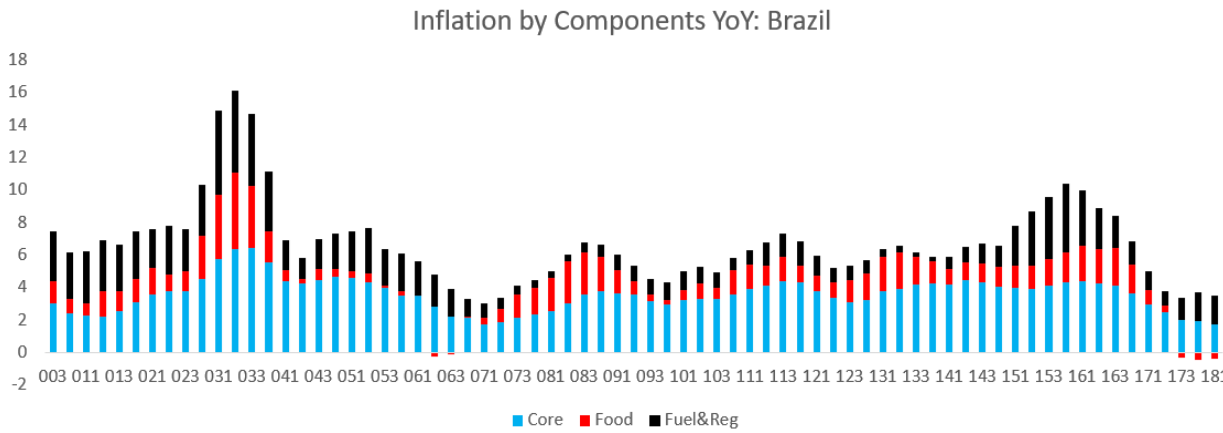
Economic cycles and inflation have a long relationship that has been captured and observed by both empirical and theoretical studies. In the past 30 years, since inflation has dropped globally to low levels, there has been increasing evidence that the effect of domestic economic slack in inflation has reduced significantly. However, in the case of Brazil, Chile, Mexico, Peru and Colombia, a first look seems to suggest that there is still some significant relationship between both (see Figure I) within the past 18 years.

Figure I



When asking these questions, the first step is to define exactly what type of inflation we are interested in studying. Typically, when policy makers and authors mention inflation they usually refer to total CPI. However, this index may not fully reflect the outcome of economic determinants, since some of its components can have seasonal effects or might be the result of direct regulations (for example, in the case of utilities or fuel). Therefore, we proceed to work with core inflation measurements, as defined by local authorities. Although this implies reducing the scope of prices of certain goods and services analyzed, it nevertheless, as can be seen in Figure II, still reflects the main driver behind total CPI inflation,.

Figure II



13 Estimation and Results

Our goal is to obtain augmented Phillip curves estimates, in order to capture and quantify the effects of both foreign and domestic factors in inflation. Formally, we proceed to estimate these for Brazil, Peru, Chile, Mexico and Colombia. The methodology used is the "general to specific" approach (GenSpec), as in Abdih et al, 2018. The first step consists of assuming a very general estimation approach, that is, estimate an equation with a predefined number of estimators, also known as the general unrestricted model (GUM). The algorithm then proceeds by deleting those estimators that are found to not be statistically significant until and reestimate again. The process continues until a final regression specification is reached. This final specification is then robust-checked by resimulating with a different deletion paths. If a final specification survives the robustness check process, it is delivered as the algorithm's final output . These final specifications are then estimated for each country, allowing us to obtain the results shown here. We then proceed to compare those results against those estimated in other papers, and also against other estimation techniques used with our own data.

Our initial GUM equation consists of six main variables: Cycle (measured by HP filter), past inflation, expectations, trade-weighted inflation of each country's 5 biggest trading partners (excluding Argentina), effective exchange rates (as defined by Haver) and oil prices (WTI). The first three account for domestic factors, whereas the latter three summarize foreign ones. Initially, since we work quarterly data, we choose the amount of lags for all variable to be 4 (one year), with the exception of inflation, since we suspect that lags effects could be longer.

$$\pi_t = c + \underbrace{\sum_{i=0}^{4} \delta_i Cycle_{t-i}}_{\text{Demand}} + \underbrace{\sum_{i=}^{9} \alpha_i \pi_{t-i} + \sum_{i=0}^{4} \theta_i Exp12_{t-i}}_{\text{Persistence and Expectations}} + \underbrace{\sum_{i=0}^{4} \beta_i MInf_{t-i} + \sum_{i=0}^{4} \gamma_i NEER_{t-i} + \sum_{i=0}^{4} \kappa_i Oil_{t-i}}_{\text{Foreign Factors}} + \epsilon_{it} \quad (10)$$

The GenSpec algorithm will then proceed to drop any non significant lagged variables, delivering a final specification. The results of these final specifications for each country are summarized in the following sections.

13.1 Data

We chose the most representative economies of LATAM according to a GDP size criteria. We left out Argentina, since there are concerns over the validity of inflation data for years 2007-2015. This resulted in a final choice of Brazil, Colombia, Mexico, Peru and Chile. For each country, CPI Core time series were obtained from Haver database as defined by local authority. We then proceeded to deseasonalize them using X-13 methodology and calculate seasonally adjusted annualized inflation rates. Alternative time series were also used as robustness checks. We then created a "foreign inflation index" that consists of an import-weighted average CPI index of the top 5 trading partners of each country (excluding Argentina). Exchange rates (NEER) were obtained from IMF database, and consist also of trade weighted exchange rate indexes. Cycles were constructed with a quarterly adjusted H-P filter. Finally, the expectations survey used are 12-months-ahead forecasts. We also used Consensus data (24 months) in robustness checks. The frequency of choice is quarterly. Length of data varies by country between 18 to 15 years, up to quarter 1 of 2018.

13.2 Main Results

We estimated the GenSpec algorithm for all five countries. Results can be found in tables I and II. Given that there several lags, result shown in the tables represent the cumulative effect (sum of all lags used) of each variable on inflation. In summary, we find that persistence is highly relevant for all countries. Regarding expectations, they play a much smaller role in all countries, with the exception of Brazil. With respect to Phillips curve slopes, we found them to be statistically significant in all cases but economically small, with the exception of Chile. These three drivers represent domestic factors. Based on our results, we can conclude that they still play a major role in determining local core inflation.

When looking at foreign factors, we also find interesting results. First and most important is that foreign inflation can impact significantly on core inflation (as much as the cycle). This is result is particularly stronger for Colombia. Second and perhaps a bit less surprising, exchange rates also might have an effect on inflation, particularly when including lags. In this sense, we also find that incorporating lags is important when estimating total effects of all variables. That is, statistical significance can be lost if these are not included in the specification. Finally, as a robustness check, we find that exchange rate and cycle estimates are more robust to alternative specification than foreign inflation ones.

Table I

	Baseline Results							Other publications(2)		
Variable	Colombia	Mexico	Peru	Chile	Brazil	Panel	Panel - WEO Specification	WEO	US(1)	EU(1)
Persistence	0.47 ***	0.73 ***	0.78 ***	0.57 ***	0.19	0.58 ***	0.64 ***	0.49 ***	0.62 ***	0.65 ***
Foreign Inflation	**0.94** ***	0.55 ***	0.37 **	0.42	0.03	**0.87** ***	0.43 *	0.02 ***	0.05 ***	-
Exchange Rate	-0.13 ***	-0.08 ***	-0.09 **	-0.07 **	0.05 ***	-0.03 **	-0.04 **	-	-	-0.02 ***
Cycle	0.26 ***	0.10 **	0.12 **	0.48 ***	0.19 **	0.14 ***	0.19 **	0.16 ***	-0.04 ***	-0.13 ***
Expectations - annual	0.47 **	-	0.20 *	0.62 *	**1.04** ***	0.49 ***	0.30	0.59 ***	0.25 ***	0.34 ***
Oil	-	-	-0.01 **	-	-	-	-	-	-	-
Time Interval	03-18	00-18	01-18	01-18	02-18	03-18	03-18	04-18	96-15	99-15
FE						X	X	X	X	X
R-squared	0.80	0.89	0.90	0.81	0.82	0.76	0.82	0.53		0.95
N	58	74	74	67	64	299	333	633	80	68

(1) Use U_Gap as cycle proxy

(2) Expectations are 24 or 36 months-ahead

*,**,*** denote p-values <0.1, <0.05 and <0.01 respectively

Table I. This table shows the result of the GenSpec algorithm applied to each country database. Dashed lines indicated that these variables did not survive the code's selection process. Numbers represent the sum of all lagged estimators for that particular variable, as well as a statistical significance test for the sum of all used estimators.

We then proceed to compare our findings with those obtained in other papers Abdih et al, 2016 and 2017. These authors use also the GenSpec approach. Additionally, we compare our findings to those obtained in the World Economic Outlook, October 2018 (WEO), where they estimate the augmented Phillips Curve for all Latin America according to the following specification:

$$\pi_t = \alpha_i \pi_{t-1} + \beta \pi_i^e + \delta Cycle_t + \beta_i MInf_t + \gamma_i NEER_t + \eta_t + \varepsilon_t \tag{11}$$

Our results are similar with these alternative findings. In all cases persistence remains as a significant factor. Also, just as in our paper, the PCs slopes that they find a statistically significant but economically small. In the case of the US and WEO-LATAM, foreign inflation also plays a significant role (units in these columns are to be read multiplied by 100), although caution is recommended when reading those estimates since foreign prices and exchange rates are collapsed together into one "foreign factors" variable.

As mentioned above, we also find that lags becomes a very important factor when determining effects of certain variables on inflation, specially when it comes to exchange rates and cycle estimates. Table II shows the case of applying WEO specification estimates to our data. The biggest difference between our baseline estimation and WEO's is the amount of lags that we incorporate. When looking at the exchange rate row, we see that all estimates loose significance once lags are dropped. Therefore, regarding future research, its important to acknowledge that many of the inflationary effects of these shocks may take some time to play a role, and such consideration must be incorporated in models. Finally, it is worth mentioning that as an additional robustness check we proceeded to reestimate these parameters by using consensus 24 months ahead expectations survey as our "expectations" variable, with similar results as in our main estimation.

Table II

Variable	WEO Type Estimation Colombia	Mexico	Peru	Chile	Brazil	Baseline Colombia	Mexico	Peru	Chile	Brazil
Persistence	0.57 ***	0.67 ***	0.79 ***	0.63 ***	0.29 ***	0.47 ***	0.73 ***	0.78 ***	0.57 ***	0.19
Foreign Inflation	0.31	0.29	-0.02	1.00 **	1.00 ***	0.94 ***	0.55 ***	0.37 **	0.42	0.03
Exchange Rate	-0.03	-0.01	-0.04	-0.04	-0.02	-0.13 ***	-0.08 ***	-0.09 **	-0.07 **	0.05 ***
Cycle	0.26 **	0.04	0.19 ***	0.19 ***	0.06	0.26 ***	0.10 **	0.12 **	0.48 ***	0.19 **
Expectations - annual	0.50 **	0.14	0.22 *	0.56	1.10 ***	0.47 **	-	0.20 *	0.62 *	1.04 ***
Time Interval	03-18	00-18	01-18	01-18	02-18	03-18	00-18	01-18	01-18	02-18
R-squared	0.69	0.92	0.87	0.75	0.81	0.80	0.89	0.90	0.81	0.82
N	56	76	73	66	60	58	74	74	67	64

*,**,*** denote p-values <0.1, <0.05 and <0.01 respectively

Table IV. This table shows the result of the GenSpec algorithm applied to each country database vs the result of using the same specification as in WEO. Dashed lines indicated that these variables did not survive the code's selection process. Numbers represent the sum of all lagged estimators for that particular variable, as well as a statistical significance test for the sum of all used estimators.

13.3 Rolling Windows

The next step is to test the hypothesis that PC slopes have "flattened". We do this in two ways. First, we reestimate the parameter associated to the cycle with 10 years rolling windows for each country. Additionally, we add as a robust check estimations of the same parameter with a Kalman Filter. The key difference is that the Kalman Filter uses all information available from $t = 0$ until t, (it never drops past information) whereas the rolling windows approach estimates only based on a fixed moving set of data throughout time. Results can be seen in Figures III.A-III.E to IX. With the exception of Mexico, results show that PC slopes have become increasingly flatter, in line with other studies.

Figure III.A

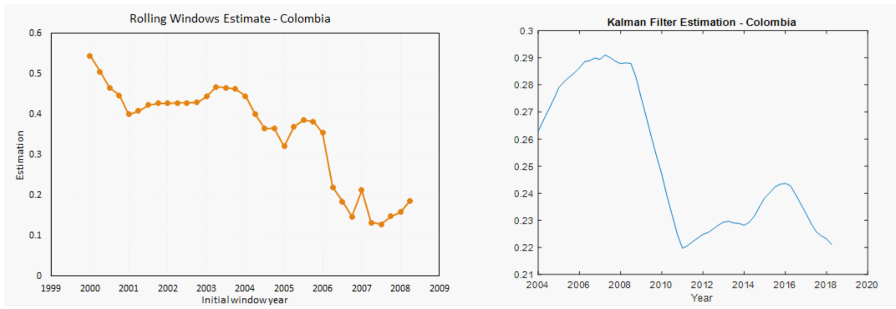


Figure III.B

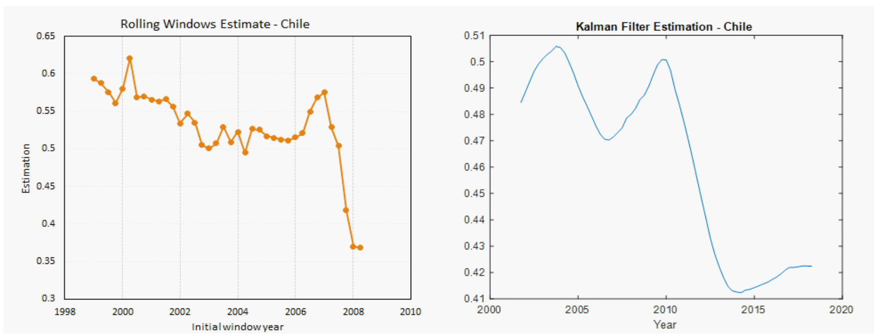


Figure III.C

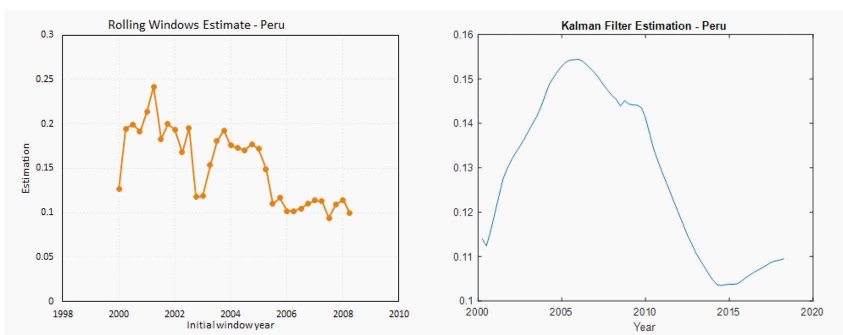


Figure III.D

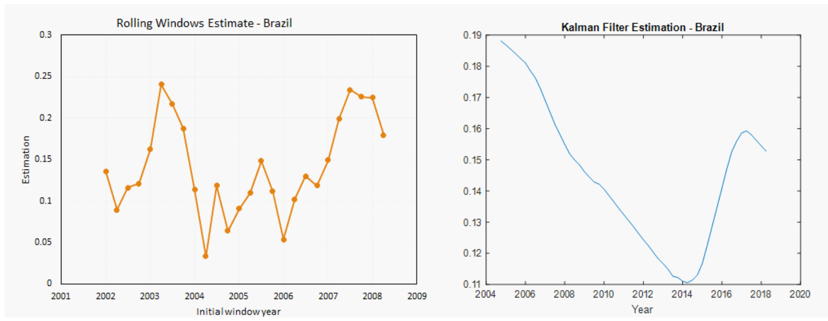
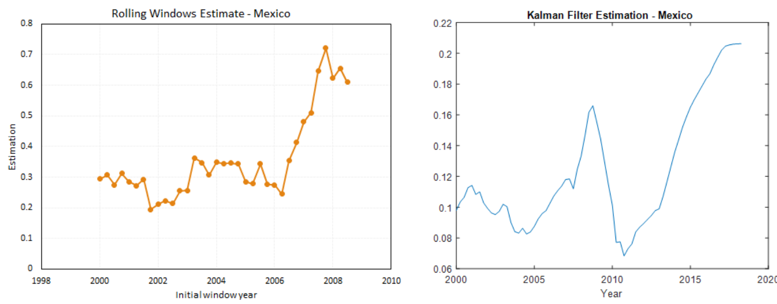


Figure III.E



13.4 Secondary results

Although it is not the intention of the paper, part of our estimations consist of measuring the effect of changes in the exchange rate on prices, also known as *pass-through.* As a robustness check, we proceed to compare them with those of another study in the Regional Economic Outlook (REO), 2016. In this study, they estimate passtrough for the same countries in our sample. Results can be contrasted in table X. We find that our results are similar to those in REO.

Table III

	Exchange Rate Pass-Through	
	Baseline	REO(1)
Colombia	0.13	0.07
Mexico	0.08	0.05
Peru	0.09	0.13
Chile	0.07	0.06
Brazil	-0.05	0.10

(1) Uses headline inflation instead

Table III. This table contrasts our pass-though estimates with those in REO, 2016

13.5 Inflation Decomposition

The previous results show statistical significance. However, it also important to check for economical significance. Therefore, we proceed to measure the economic impact of each variable as follows. First, we redefine inflation (and inflation expectations) as deviations from central bank's target. This allows us to compare more easily between countries. We then quantify the contribution of each factor to each quarter's inflation by multiplying the change in variables and the estimated parameters, as in the following equation:

$$\pi_t - \bar{\pi}_i = \underbrace{\hat{\alpha}_1(\pi_{it-1} - \bar{\pi}_i) + \hat{\alpha}_2(\pi_{it-2} - \bar{\pi}_i)}_{\text{Contribution of persistence to deviation}} + \underbrace{\hat{\theta}\,(Exp12_{it-1} - \bar{\pi}_i)}_{\text{Contribution of expectations to deviation}} + \underbrace{\beta X_i}_{\text{Contribution of other explanatory variables}} + \underbrace{(1 - \hat{\alpha}_1 - \hat{\alpha}_2 - \hat{\theta})\bar{\pi}_i + \hat{c} + \epsilon_{it}}_{\text{Residual}} \qquad (12)$$

This calculation allows us to obtain the contribution during each quarter for each variables in inflation. Therefore, a high contribution may be the result of two things: high statistical significance in the estimator associated to the variable and a high level of volatility in the variable itself. Figures IV.A-IV.E show contribution per variables by adding up all the lagged effects for each of them.

Figure IV.A

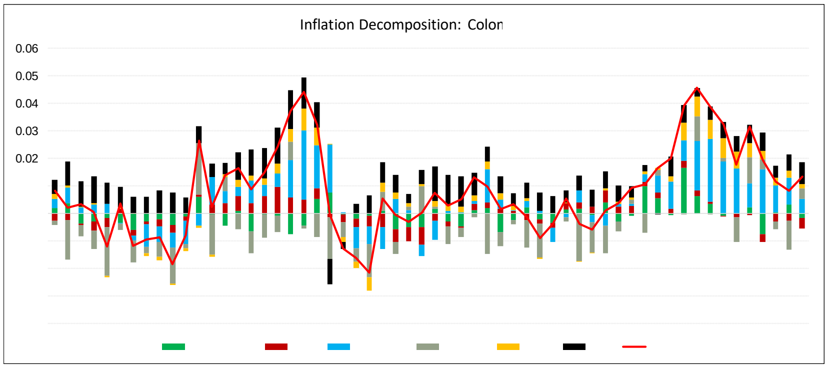


Figure IV.B

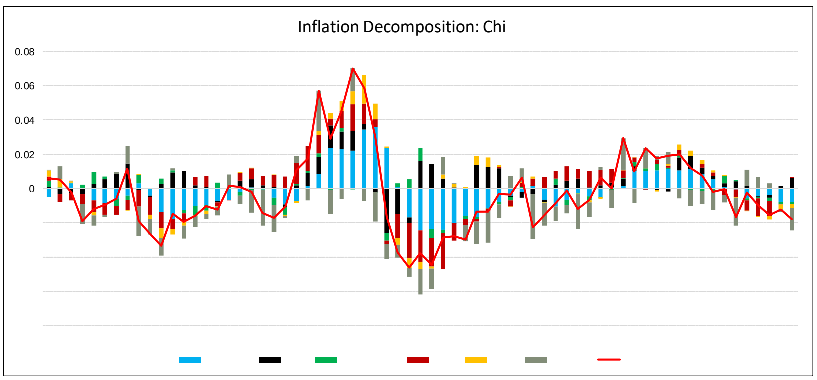


Figure IV.C

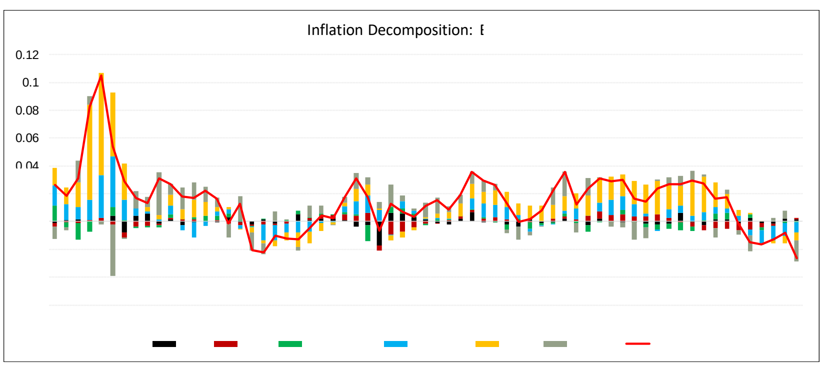


Figure IV.D

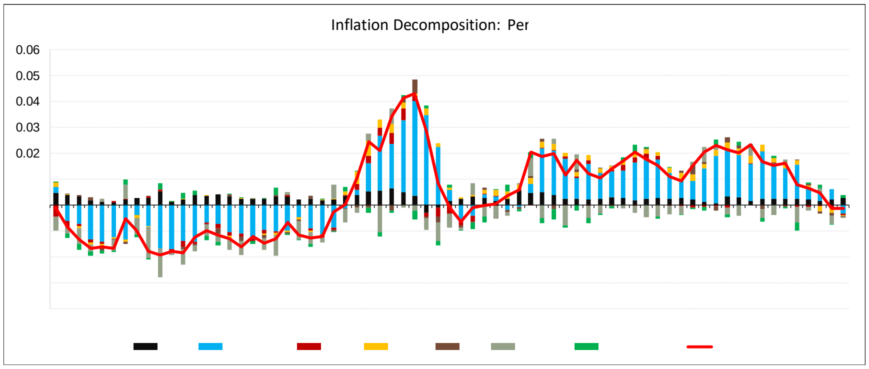
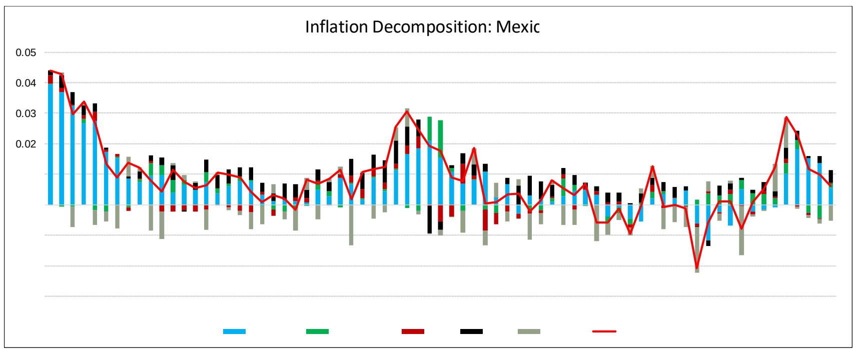


Figure IV..E



Figures IV.A-IV.E. Colored columns denote the economic impact of each variable in inflation points as deviation from inflation target. Minf represents foreign inflation and Exp12 represent 12 month ahead expectations as relevated by local central bank.

The previous figures show the effects for some countries. As can be seen, there are certain quarters where some factors may play a bigger role than others. However, a more holistic view may be also of interest. Figure V shows total effects per variable as a fraction of total inflation deviation from target. It was constructed the following way: we measure each effect per variable per quarter (as in Figures IV.A-IV.E), and then proceed to add them up for all the time interval (in absolute values, to avoid negative shocks compensating for positive ones). Then, we measure each of them proportionally to total shocks on inflation. That is, Figure V shows the total proportion of each variable in inflationary shocks during the whole sample period for each country. In summary, the main findings are several. First of all, domestic factor still a big role. Persistence, expectations and cycle still explain at least 50% of deviations from inflation targets, if not more. Particularly, we find that persistence is still the most relevant among them. Second, foreign inflation can play a significant role, especially in some countries like Colombia or Brazil. Foreign inflation can shock inflation as much as expectations or persistence, in some cases. Third, domestic slack matters, but by less than many other factors. Though it still plays a role in determining inflation, it is not the main driver by itself. Even in some countries like Peru and Brazil it is among the lease relevant factors.

Figure V

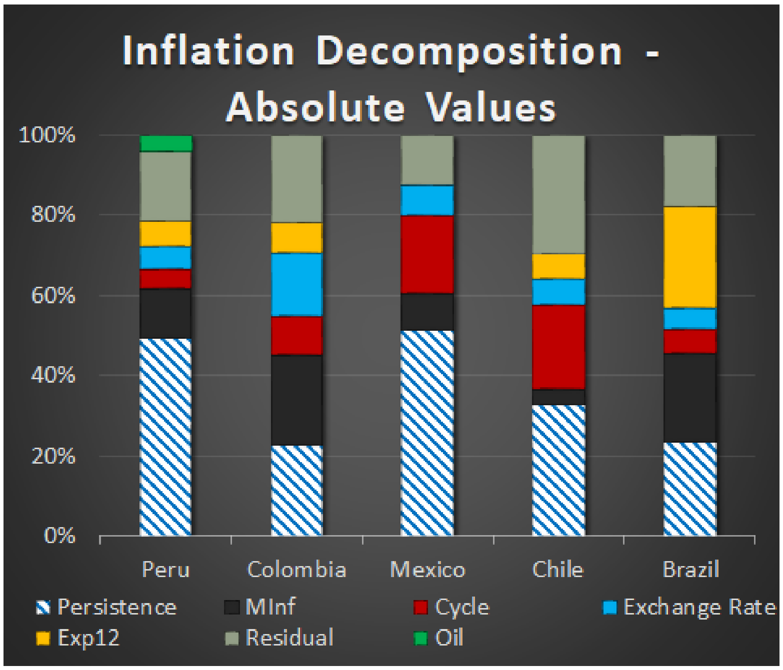


Figure V. Colored columns denote the economic impact of each variable in inflation points as a fraction of the total inflationary impact (in absolute values). Minf represents foreign inflation and Exp12 represent 12 month ahead expectations as relevated by local central bank.

13.6 Additional estimations

As an additional experiment, we ran the same algorithm but restricting on tradables and non-tradables. Results can be seen Tables IV and V. For Tradables we find that, not surprisingly, foreign inflation can play an important role in determining their inflation. Also, for both Tradables and Non-Tradables we find that, as expected, either persistence and/or expectations is still among the most relevant factors. Influence of cycle is less straightforward and varies depending on each country.

Finally, we also tested for nonlinearity in the PC, as in Abdih, Balakrishnan and Shang (2016), finding non-statistically significant results for all countries. That is, we do not find evidence that PC slopes might have a non-linear shape, similar to what the same authors find in their paper for the US.

Table IV

Summary Tradables						
Variable	Colombia	Mexico	Peru	Chile	Brazil	Panel
Persistence	0.714565 ***	0.759213 ***	-0.299021 **	-	0.884404 ***	0.595568 ***
Foreign Inflation	0.999547 **	0.428352 **	0.387667	2.851727 ***	1.155112 ***	1.963529 ***
Exchange Rate	-0.385851 ***	-	0.031604	-	-	-0.077017 ***
Cycle	-	-	0.4755 ***	0.765072 ***	0.102238 **	-
Expectations - annual	-	-	-	-	-	-
Oil	-	-	-	-	-	-
R-squared	0.8	0.68	0.61	0.5	0.97	0.45
N	65	64	62	70	54	319

*,**,*** denote p-values <0.1, <0.05 and <0.01 respectively

Table V

Summary Non Tradables						
Variable	Colombia	Mexico	Peru	Chile	Brazil	Panel
Persistence	0.5227276 ***	0.0003383	-0.3094402 *	-	0.8394881	0.3852347 ***
Foreign Inflation	-	2.0185173 **	-1.364127 **	-	0.3706982 **	1.132785 ***
Exchange Rate	-	-0.1083204 **	-	-0.3983998 ***	-0.0242632 **	-0.0550736 **
Cycle	0.0471281	0.0308082	0.8134532 ***	-	0.0567349	0.091016
Expectations - annual	0.28061 ***	-	-	2.95373 ***	-0.12505 **	-
Oil	-	-	-	-	-	-
R-squared	0.61	0.72	0.45	0.41	0.96	0.41
N	58	62	62	67	54	274

*,**,*** denote p-values <0.1, <0.05 and <0.01 respectively

Tables IV and V. These tables show the result of the GenSpec algorithm applied to each country database, restricting for only tradable and non-tradable goods, respectively. Dashed lines indicated that these variables did not survive the code's selection process. Numbers represent the sum of all lagged estimators for that particular variable, as well as a statistical significance test for the sum of all used estimators.

14 Conclusion

Our paper intends to estimate augmented Phillips curves for five major LATAM countries. We do this by using a General-to-Specific approach, an algorithm already applied in other studies that focus in other economic regions, such as the EU or USA. This algorithms consists in going from a very general PC regression specification to a more specific one in a systematic way, in which non-statistically significant variables are dropped fro the specification. We apply it to Brazil, Mexico, Chile, Peru and Colombia, for the last 18 years. Our main goal is to estimate what are the main drivers of inflation in these countries.

In summary, we find that domestic factors still play a major role in determining inflation. Past inflation and expectations still seem to be the main drivers behind inflationary processes, with economic slack playing still a role but smaller. Persistence is overall the most important variable. Expectations on the other hand vary depending on country and at what horizon they are. Therefore expectations formation (either backward or forward) becomes key in understanding inflation processes better. Also, cycles matter, but not more than other factors. Foreign drivers like foreign inflation of major trade partners can actually be more relevant, specially for Colombia. Regarding the change of PC slopes throughout time, we find that they have reduced for most LATAM countries. In this sense, our findings are similar to others done for EMs and Developing economies. We also find that exchange rate pass through estimates also similar to REO 2016 findings. Finally, we also our results and robust check them with other methodologies used in other studies like WEO, finding similar results.

15 References

Abdih, Y., Balakrishnan, R., Shang, B., 2016, "What is Keeping U.S. Core Inflation Low: Insights from a Bottom-up Approach," IMF Working Paper No. 16/124.

Abdih, Y., Lin L., Paret, A., 2018, "Understanding Euro Area Inflation Dynamics: Why So Low for So Long?", IMF Working Paper No 18/188.

Auer, R., Borio, C., Filardo, A., 2017, The Globalisation of Inflation: The Growing Importance of Global Value Chains. Bank for International Settlements Working Papers, No. 602.

International Monetary Fund, 2018, "Determinants of Inflation in Emerging Markets", World Economic Outlook, October 2018, Chapter 3.

International Monetary Fund, 2016, "Exchange Rate Pass-Through in Latin America", Regional Economic Outlook Western Hemisphere, April 2016, Chapter 4.

Gali J., Gertler, M., 1999, "Inflation Dynamics: A Structural Econometric Analysis," Journal of Monetary Economics, Volume 44, Issue 2, p. 195.

Lanau, S., Robles, A., Toscani, F., 2018, "Explaining Inflation in Colombia: A Disaggregated Phillips Curve Approach", IMF Working Paper No 18/106.

Naudon, Alberto and Vial, Joaquín, "The Evolution of Inflation in Chile Since 2000",. BIS Paper, November 2016, No. 89g.

Part IV

Appendixes

16 Appendix Part 1

16.1 *Additional Figures*

Figure III

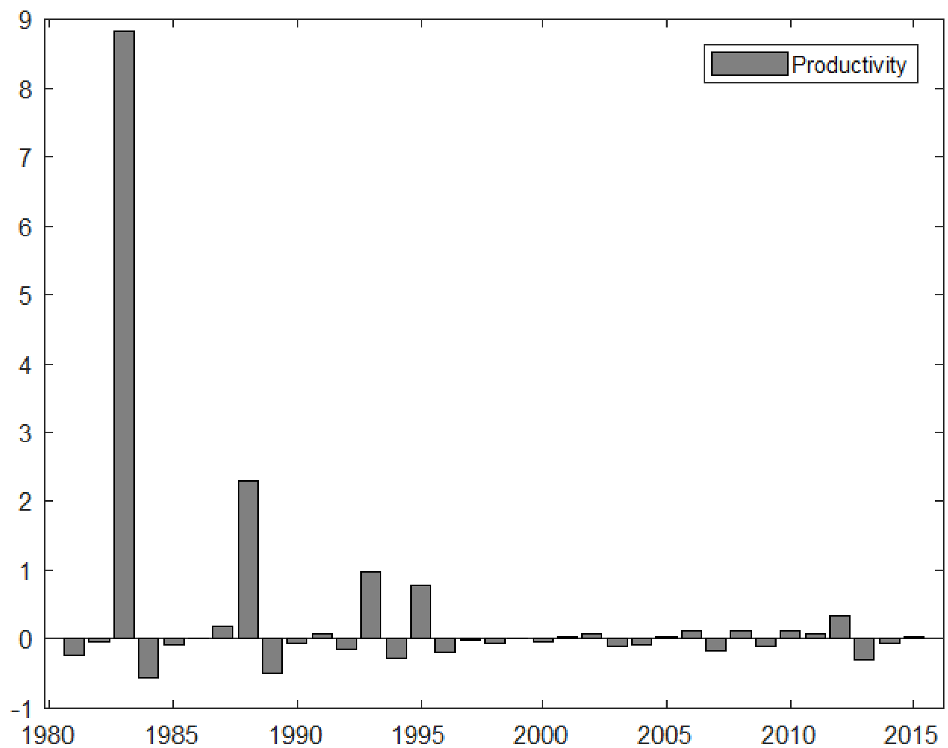


Figure III. Fig. III shows the contribution of each factor in total price change (measured in percentage points change with respect to the previous year). More specifically, Fig. III shows changes in prices due to productivity shocks. They were calculated by replacing the estimated values of Z_t^i, A_t,ω_i and η , $i \in$ {food,feed,ethanol,exports} , into equation (2). Each term in equation (2) is represented by a different bar for a given year.

Figure V

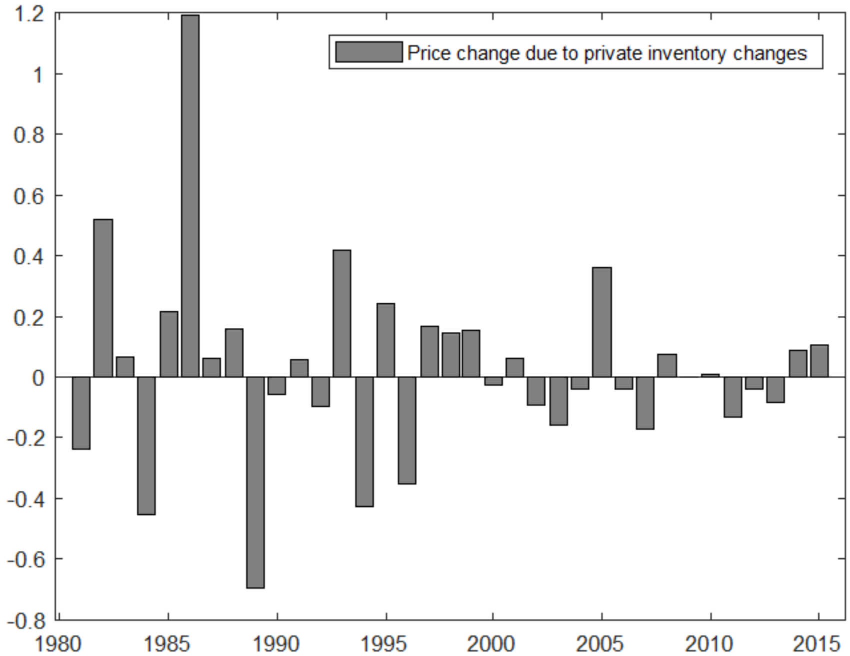


Figure V. Fig. V shows the contribution of total price change due to private inventory purchases. It was calculated by estimating the change in prices not explained by equation (2), i.e., the residual between explained price changes and observed price changes. That residual was then multiplied by the proportion of corn inventories not held by the government under the CCC program.

Figure VI

Price Indexes for different Commodities

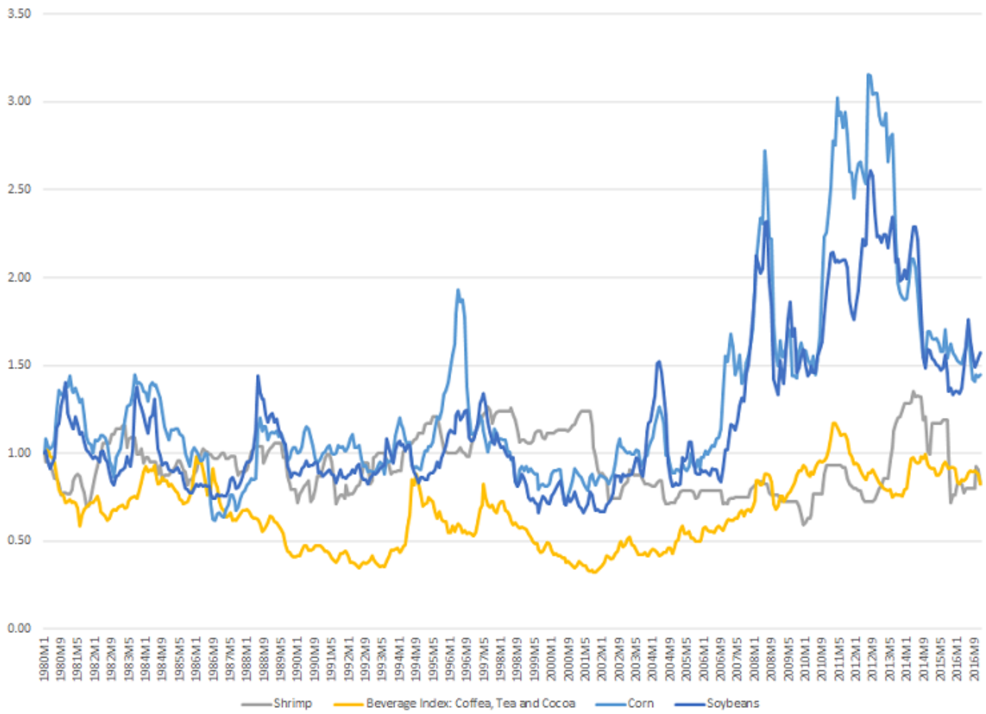
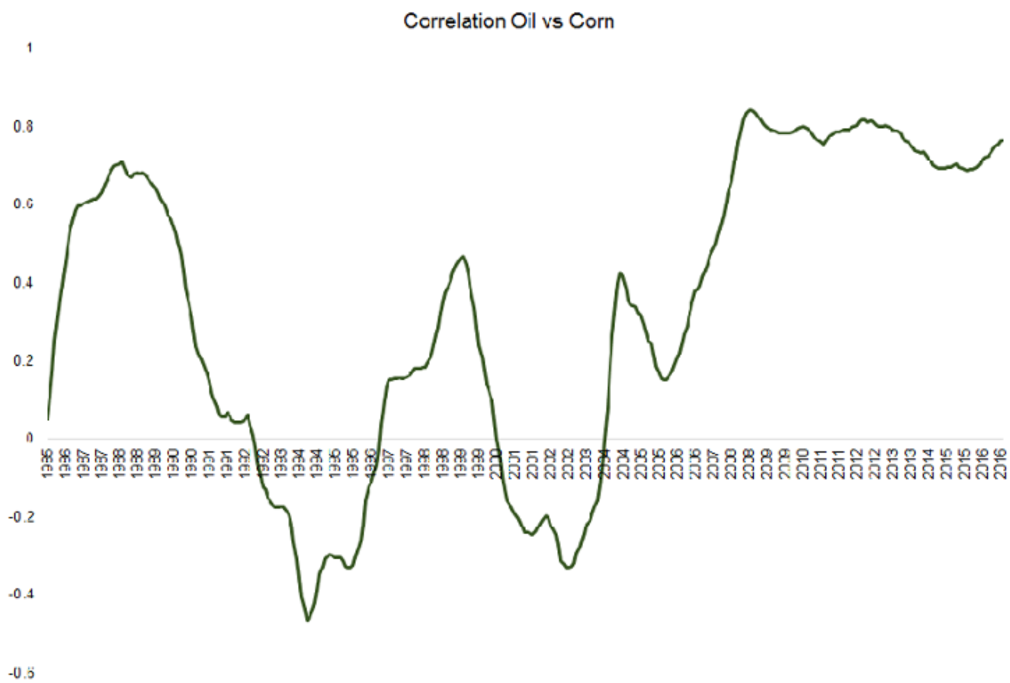


Figure VII



Figures VI-VII. Fig. VI shows how corn and soybean prices have deviated from other commodities that were not used for energy production. Fig. VII shows increasing daily correlation between corn prices and oil prices for a past-five-years rolling window.

Figure XII

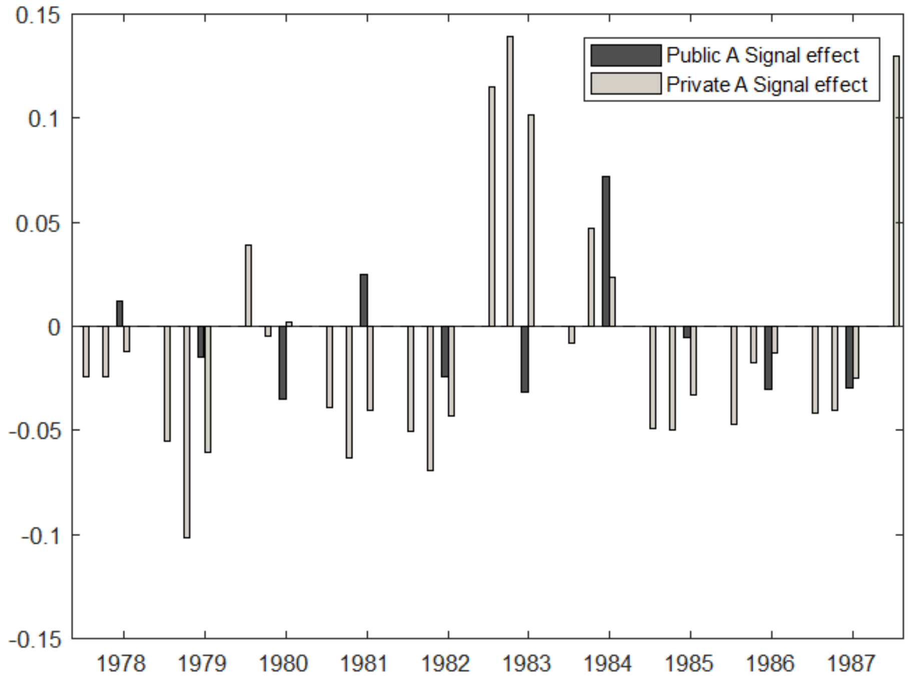


Figure XII. Fig. XII shows the effect on prices for forecasts regarding supply. The effects are shown as changes in prices in percentage points.

Figure XIII

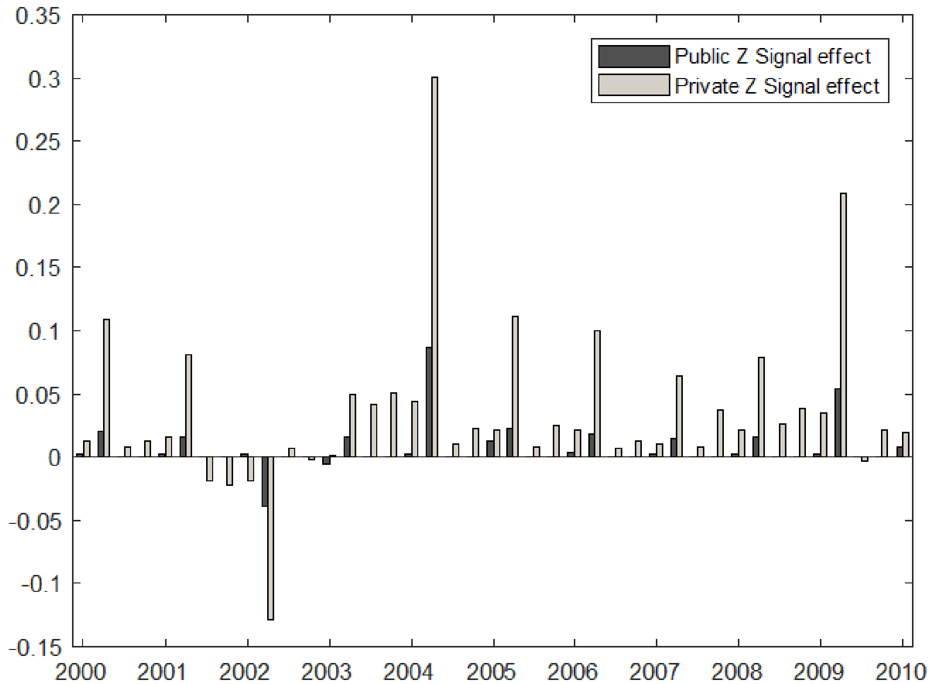


Figure XIII. Fig. XIII shows the effect on prices for forecasts regarding demand. The effects are shown as changes in prices in percentage points.

16.2 High-Frequency Model.Proof of parameter estimates

Parameters that associate state variables with prices were calculated using the method of undetermined coefficients. Below are the steps and formal results of such calculations for each quarter. The solving pattern is identical in all four quarters; therefore, I mainly describe the first quarter and solve the following ones in the same way. Since final expressions are implicit, solutions were found numerically.

16.2.1 *First quarter*

Expectations equation in quarter 1 results:

$$E_1(p_2) = \alpha_1^{II} E_1\left[\tilde{A}_1\right] + \alpha_2^{II} E_1\left[\varepsilon_2^Z\right] + \alpha_3^{II} E_1\left[X_2\right] + \alpha_4^{II} E_1\left[\theta_2^A\right] + \alpha_5^{II} E_1\left[\theta_2^Z\right] + \alpha_8^{II} E_1\left[Z_y\right].$$

Replacing the speculator's policy function in the previous equation:

$$\delta 2 X_2 + \tfrac{p_1}{\beta} = \alpha_1^{II} E_1\left[\tilde{A}_1\right] + \alpha_2^{II} E_1\left[\varepsilon_2^Z\right] + \alpha_3^{II} E_1\left[X_2\right] + \alpha_4^{II} E_1\left[\theta_2^A\right] + \alpha_5^{II} E_1\left[\theta_2^Z\right] + \alpha_8^{II} E_1\left[Z_y\right].$$

$$\tfrac{p_1}{\beta} = \alpha_1^{II} E_1\left[\tilde{A}_1\right] + \alpha_2^{II} E_1\left[\varepsilon_2^Z\right] + \left(\alpha_3^{II} - 2\delta\right) X_2 + \alpha_4^{II} E_1\left[\theta_2^A\right] + \alpha_5^{II} E_1\left[\theta_2^Z\right] + \alpha_8^{II} E_1\left[Z_y\right].$$

I now replace X_2 with the market-clearing equation for the first quarter and leave the price variable on the left-hand side so to match the linear solution equation:

$$\tfrac{p_1}{\beta} = \begin{bmatrix} \alpha_1^{II} E_1\left[\tilde{A}_1\right] + \alpha_2^{II} E_1\left[\varepsilon_2^Z\right] \\ + \left(\alpha_3^{II} - 2\delta\right) \begin{bmatrix} \bar{p}_1^{\eta} \tilde{A}_1 - \bar{p}_1^{-\rho}\left[Z_y + \varepsilon_1^z\right] + X_1 \\ +(\eta \bar{A} \bar{p}_1^{\eta-1} + \rho \bar{Z}_1' \bar{p}_1^{-\rho-1}) p_1 \end{bmatrix} + \\ +\alpha_4^{II} E_1\left[\theta_2^A\right] + \alpha_5^{II} E_1\left[\theta_2^Z\right] + \alpha_8^{II} E_1\left[Z_y\right] \end{bmatrix}.$$

$$p_1\left[\tfrac{1}{\beta} - \left(\alpha_3^{II} - 2\delta\right)(\eta \bar{A} \bar{p}_1^{\eta-1} + \rho \bar{Z}_1' \bar{p}_1^{-\rho-1})\right] = \begin{bmatrix} \alpha_1^{II}\left[\theta_1^A + \rho_A A_0\right] \\ + \left(\alpha_3^{II} - \delta 2\right)\left[\bar{p}_1^{\eta} \tilde{A}_1 - \bar{p}_1^{-\rho}\left[Z_y + \varepsilon_1^z\right] + X_1\right] + \\ +\alpha_4^{II} E_1\left[\theta_2^A\right] + \alpha_5^{II} E_1\left[\theta_2^Z\right] \\ +\alpha_8^{II}\left[\rho_z Z_{y-1} + \theta_1^Z\right] \end{bmatrix}.$$

$$p_1\left[\tfrac{1}{\beta} - \left(\alpha_3^{II} - 2\delta\right)(\eta \bar{A} \bar{p}_1^{\eta-1} + \rho \bar{Z}_1' \bar{p}_1^{-\rho-1})\right] = \begin{bmatrix} \alpha_1^{II}\theta_1^A + \alpha_1^{II}\rho_A A_0 + \alpha_8^{II}\theta_1^Z + \beta_8 \rho_z Z_{y-1} \\ -\left(\alpha_3^{II} - 2\delta\right)\bar{p}_1^{-\rho} Z_y + + \left(\alpha_3^{II} - 2\delta\right)\bar{p}_1^{\eta}\tilde{A}_1 \\ -\left(\alpha_3^{II} - 2\delta\right)\bar{p}_1^{-\rho}\varepsilon_1^z + \left(\alpha_3^{II} - 2\delta\right) X_1 + \\ +\alpha_4^{II} E_1\left[\theta_2^A\right] + \alpha_5^{II} E_1\left[\theta_2^Z\right] \end{bmatrix}.$$

$$p_1\left[\tfrac{1}{\beta} - \left(\alpha_3^{II} - 2\delta\right)(\eta \bar{A} \bar{p}_1^{\eta-1} + \rho \bar{Z}_1' \bar{p}_1^{-\rho-1})\right] = \begin{bmatrix} \left(\alpha_3^{II} - 2\delta\right)\bar{p}_1^{\eta}\tilde{A}_1 - \left(\alpha_3^{II} - 2\delta\right)\bar{p}_1^{-\rho}\varepsilon_1^z \\ + \left(\alpha_3^{II} - 2\delta\right) X_1 \\ +\alpha_1^{II}\theta_1^A + \alpha_2^{II}\theta_1^Z + \alpha_1^{II}\rho_A A_0 \\ +\alpha_8^{II}\rho_z Z_{y-1} - \left(\alpha_3^{II} - 2\delta\right)\bar{p}_1^{-\rho} Z_y \end{bmatrix}.$$

Therefore, the final price solution equation is:

$$
\begin{aligned}
p_1 \;=\;& \frac{\left(\alpha_3^{II}-2\delta\right)\bar{p}_1^{\eta}}{\left[\frac{1}{\beta}-\left(\alpha_3^{II}-2\delta\right)\left(\eta\bar{A}\bar{p}_1^{\eta-1}+\rho\bar{Z}_1'\bar{p}_1^{-\rho-1}\right)\right]}\tilde{A}_1-\frac{\left(\alpha_3^{II}-2\delta\right)\bar{p}_1^{-\rho}}{\left[\frac{1}{\beta}-\left(\alpha_3^{II}-2\delta\right)\left(\eta\bar{A}\bar{p}_1^{\eta-1}+\rho\bar{Z}_1'\bar{p}_1^{-\rho-1}\right)\right]}\varepsilon_1^{z}\\
&+\frac{\left(\alpha_3^{II}-2\delta\right)}{\left[\frac{1}{\beta}-\left(\alpha_3^{II}-2\delta\right)\left(\eta\bar{A}\bar{p}_1^{\eta-1}+\rho\bar{Z}_1'\bar{p}_1^{-\rho-1}\right)\right]}X_1+\frac{\alpha_1^{II}}{\left[\frac{1}{\beta}-\left(\alpha_3^{II}-2\delta\right)\left(\eta\bar{A}\bar{p}_1^{\eta-1}+\rho\bar{Z}_1'\bar{p}_1^{-\rho-1}\right)\right]}\theta_1^{A}\\
&+\frac{\alpha_2^{II}}{\left[\frac{1}{\beta}-\left(\alpha_3^{II}-2\delta\right)\left(\eta\bar{A}\bar{p}_1^{\eta-1}+\rho\bar{Z}_1'\bar{p}_1^{-\rho-1}\right)\right]}\theta_1^{Z}+\frac{\alpha_8^{II}\rho_z}{\left[\frac{1}{\beta}-\left(\alpha_3^{II}-2\delta\right)\left(\eta\bar{A}\bar{p}_1^{\eta-1}+\rho\bar{Z}_1'\bar{p}_1^{-\rho-1}\right)\right]}Z_{y-1}\\
&+\frac{\alpha_1^{II}\rho_A}{\left[\frac{1}{\beta}-\left(\alpha_3^{II}-2\delta\right)\left(\eta\bar{A}\bar{p}_1^{\eta-1}+\rho\bar{Z}_1'\bar{p}_1^{-\rho-1}\right)\right]}A_0-\frac{\left(\alpha_3^{II}-2\delta\right)\bar{p}_1^{-\rho}}{\left[\frac{1}{\beta}-\left(\alpha_3^{II}-2\delta\right)\left(\eta\bar{A}\bar{p}_1^{\eta-1}+\rho\bar{Z}_1'\bar{p}_1^{-\rho-1}\right)\right]}Z_y.
\end{aligned}
$$

And parameter values are given by:

$$\alpha_1^I = \frac{\left(\alpha_3^{II}-\delta 2\right)\bar{p}_1^{\eta}}{\left[\frac{1}{\beta}-(\beta_3-\delta 2)(\eta\bar{A}\bar{p}_1^{\eta-1}+\rho\bar{Z}_1'\bar{p}_1^{-\rho-1})\right]}.$$

$$\alpha_2^I = -\frac{\left(\alpha_3^{II}-2\delta\right)\bar{p}_1^{-\rho}}{\left[\frac{1}{\beta}-\left(\alpha_3^{II}-\delta 2\right)(\eta\bar{A}\bar{p}_1^{\eta-1}+\rho\bar{Z}_1'\bar{p}_1^{-\rho-1})\right]}.$$

$$\alpha_3^I = \frac{\left(\alpha_3^{II}-2\delta\right)}{\left[\frac{1}{\beta}-\left(\alpha_3^{II}-\delta 2\right)(\eta\bar{A}\bar{p}_1^{\eta-1}+\rho\bar{Z}\bar{p}_1^{-\rho-1})\right]}.$$

$$\alpha_4^I = \frac{\alpha_1^{II}}{\left[\frac{1}{\beta}-\left(\alpha_3^{II}-\delta 2\right)(\eta\bar{A}\bar{p}_1^{\eta-1}+\rho\bar{Z}_1'\bar{p}_1^{-\rho-1})\right]}.$$

$$\alpha_5^I = \frac{\alpha_2^{II}}{\left[\frac{1}{\beta}-\left(\alpha_3^{II}-\delta 2\right)(\eta\bar{A}\bar{p}_1^{\eta-1}+\rho\bar{Z}_1'\bar{p}_1^{-\rho-1})\right]}.$$

$$\alpha_6^I = \frac{\alpha_8^{II}\rho_z}{\left[\frac{1}{\beta}-\left(\alpha_3^{II}-\delta 2\right)(\eta\bar{A}\bar{p}_1^{\eta-1}+\rho\bar{Z}_1'\bar{p}_1^{-\rho-1}))\right]}.$$

$$\alpha_7^I = \frac{\alpha_1^{II}\rho_A}{\left[\frac{1}{\beta}-\left(\alpha_3^{II}-\delta 2\right)(\eta\bar{A}\bar{p}_1^{\eta-1}+\rho\bar{Z}_1'\bar{p}_1^{-\rho-1})\right]}.$$

$$\alpha_8^I = -\frac{\left(\alpha_3^{II}-2\delta\right)\bar{p}_1^{-\rho}}{\left[\frac{1}{\beta}-\left(\alpha_3^{II}-\delta 2\right)(\eta\bar{A}\bar{p}_1^{\eta-1}+\rho\bar{Z}_1'\bar{p}_1^{-\rho-1})\right]}.$$

16.2.2 *Second quarter*

The expectations equation is:

$$E_2(p_3) = \alpha_1^{III}\tilde{A}_1 + \alpha_2^{III}E_2\left[\varepsilon_3^{Z}\right] + \alpha_3^{III}E_2\left[X_3\right] + \alpha_4^{III}E_2\left[\theta_3^{A}\right] + \alpha_5^{III}E_2\left[\theta_3^{Z}\right] + \alpha_8^{III}E_2\left[Z_y\right].$$

The solution equation and the speculator's policy function are:

$\delta^{II}2X_3 + \frac{p_2}{\beta} = \alpha_1^{III}\tilde{A}_1 + \alpha_3^{III}X_3 + \alpha_4^{III}\theta_2^A + \alpha_5^{III}\theta_2^Z + \alpha_8^{III}Z_y.$

$\frac{p_2}{\beta} = \alpha_1^{III}\tilde{A}_1 + \left(\alpha_3^{III} - \delta^{II}2\right)X_3 + \alpha_4^{III}\theta_2^A + \alpha_5^{III}\theta_2^Z + \alpha_8^{III}Z_y.$

I introduce the market-clearing equation and solve:

$\frac{p_2}{\beta} = \alpha_1^{III}\tilde{A}_1 + \left(\alpha_3^{III} - \delta^{II}2\right)\left[-\bar{p}_2^{-\rho}\tilde{Z}_2 + \rho\bar{Z}\bar{p}_2^{-\rho-1}p_2 + X_2\right] + \alpha_4^{III}\theta_2^A + \alpha_5^{III}\theta_2^Z + \alpha_8^{III}Z_y.$

$$p_2\left[\frac{1}{\beta} - \rho\bar{Z}_2\bar{p}_2^{-\rho-1}\left(\alpha_3^{III} - \delta^{II}2\right)\right] = \left[\begin{array}{c} \alpha_1^{III}\tilde{A}_1 - \left(\alpha_3^{III} - \delta^{II}2\right)\bar{p}_2^{-\rho}\left[Z_y + \varepsilon_z^2\right] \\ +\left(\alpha_3^{III} - \delta^{II}2\right)X_2 + \\ +\alpha_4^{III}\theta_2^A + \alpha_5^{III}\theta_2^Z + \alpha_8^{III}Z_y \end{array}\right].$$

$$p_2\left[\frac{1}{\beta} - \rho\bar{Z}_2\bar{p}_2^{-\rho-1}\left(\alpha_3^{III} - \delta^{II}2\right)\right] = \left[\begin{array}{c} \alpha_1^{III}\tilde{A}_1 - \left(\alpha_3^{III} - \delta^{II}2\right)\bar{p}_2^{-\rho}\varepsilon_z^2 \\ +\left(\alpha_3^{III} - \delta^{II}2\right)X_2 + \alpha_4^{III}\theta_2^A + \\ +\alpha_5^{III}\theta_2^Z + \left(\alpha_8^{III} - \left(\alpha_3^{III} - \delta^{II}2\right)\bar{p}_2^{-\rho}\right)Z_y \end{array}\right].$$

Final price solution equation for quarter 2 is:

$$p_2 = \left[\begin{array}{c} \frac{\alpha_1^{III}}{\left[\frac{1}{\beta} - \rho\bar{Z}_2\bar{p}_2^{-\rho-1}\left(\alpha_3^{III} - \delta 2\right)\right]}A_1 - \frac{\left(\alpha_3^{III} - \delta 2\right)\bar{p}_2^{-\rho}}{\left[\frac{1}{\beta} - \rho\bar{Z}_2\bar{p}_2^{-\rho-1}\left(\alpha_3^{III} - \delta 2\right)\right]}\varepsilon_z^2 \\ +\frac{\left(\alpha_3^{III} - \delta 2\right)}{\left[\frac{1}{\beta} - \rho\bar{Z}_2\bar{p}_2^{-\rho-1}\left(\alpha_3^{III} - \delta^{II}2\right)\right]}X_2 + +\frac{\alpha_4^{III}}{\left[\frac{1}{\beta} - \rho\bar{Z}_2\bar{p}_2^{-\rho-1}\left(\alpha_3^{III} - \delta 2\right)\right]}\theta_2^A \\ +\frac{\alpha_5^{III}}{\left[\frac{1}{\beta} - \rho\bar{Z}_2\bar{p}_2^{-\rho-1}\left(\alpha_3^{III} - \delta^{II}2\right)\right]}\theta_2^Z + \frac{\left(\alpha_8^{III} - \left(\alpha_3^{III} - \delta\lambda 2\right)\bar{p}_2^{-\rho}\right)}{\left[\frac{1}{\beta} - \rho\bar{Z}_2\bar{p}_2^{-\rho-1}\left(\alpha_3^{III} - \delta 2\right)\right]}Z_y. \end{array}\right]$$

Parameter solutions are:

$\alpha_1^{II} = \frac{\alpha_1^{III}}{\left[\frac{1}{\beta} - \rho\bar{Z}_2\bar{p}_2^{-\rho-1}\left(\alpha_3^{III} - \delta 2\right)\right]}.$

$\alpha_2^{II} = -\frac{\left(\alpha_3^{III} - \delta 2\right)\bar{p}_2^{-\rho}}{\left[\frac{1}{\beta} - \rho\bar{Z}_2\bar{p}_2^{-\rho-1}\left(\alpha_3^{III} - \delta 2\right)\right]}.$

$\alpha_3^{II} = \frac{\left(\alpha_3^{III} - \delta 2\right)}{\left[\frac{1}{\beta} - \rho\bar{Z}_2\bar{p}_2^{-\rho-1}\left(\alpha_3^{III} - \delta^{II}2\right)\right]}.$

$$\alpha_4^{II} = \frac{\alpha_4^{III}}{\left[\frac{1}{\beta} - \rho \bar{Z}_2 \bar{p}_2^{-\rho-1}\left(\alpha_3^{III} - \delta 2\right)\right]}.$$

$$\alpha_5^{II} = \frac{\alpha_5^{III}}{\left[\frac{1}{\beta} - \rho \bar{Z}_2 \bar{p}_2^{-\rho-1}\left(\alpha_3^{III} - \delta^{II} 2\right)\right]}.$$

$$\alpha_8^{II} = \frac{\left(\alpha_8^{III} - \left(\alpha_3^{III} - \delta\lambda 2\right)\bar{p}_2^{-\rho}\right)}{\left[\frac{1}{\beta} - \rho \bar{Z}_2 \bar{p}_2^{-\rho-1}\left(\alpha_3^{III} - \delta 2\right)\right]}.$$

Third quarter is identical to the second one

16.2.3 *Fourth Quarter*

Solution equation and expectations equation are respectively:

$$\tilde{p}_4 = \alpha_1^{IV} \tilde{A}_1 + \alpha_2^{IV} \tilde{Z}_4 + \alpha_3^{IV} X_4 + \alpha_4^{IV} \theta_4^A + \alpha_5^{IV} \theta_4^Z + \alpha_8^{IV} Z_y.$$

$$E_4\left[\tilde{p}_{1B}\right] = \begin{bmatrix} \alpha_1^I E_4\left[\tilde{A}_{1B}\right] + \alpha_2^I E_4\left[\varepsilon_{1B}^z\right] + \alpha_3^I E_4\left[X_{1B}\right] \\ +\alpha_4^I E_4\left[\theta_{1B}^A\right] + \alpha_5^I E_4\left[\theta_{1B}^Z\right] + \alpha_6 E_4^I\left[Z_{y-1}\right] \\ +\alpha_7^I E_4\left[\tilde{A}_1\right] + +\alpha_8^I E_4\left[Z_{yB}\right] \end{bmatrix}.$$

Introducing the speculator's policy function into the expectations equation I get the following:

$$\delta^{II} 2 X_{1B} + \frac{p_4}{\beta} = \begin{bmatrix} \alpha_1^I \underbrace{E_4\left[\tilde{A}_{1B}\right]}_{\rho_A \tilde{A}_1 + \underbrace{E_4\left[\varepsilon_{1B}^A\right]}_{\theta_4^A}} + \alpha_2^I \underbrace{E_4\left[\varepsilon_{1B}^z\right]}_{=0} + \alpha_3^I E_4\left[X_{1B}\right] \\ + \underbrace{\alpha_4^I E_4\left[\theta_{1B}^A\right] + \alpha_5^I E_4\left[\theta_{1B}^Z\right]}_{=0} + \alpha_6^I Z_y + \alpha_7^I \tilde{A}_1 + \\ +\alpha_8^I \underbrace{E_4\left[Z_{yB}\right]}_{\rho_z Z_y + \underbrace{E_4\left[\varepsilon_{1B}^z\right]}_{\theta_4^Z}} \end{bmatrix}.$$

$$\delta^{II} 2 X_{1B} + \frac{p_4}{\beta} = \left(\alpha_1^I \rho_A + \alpha_7^I\right) \tilde{A}_1 + \alpha_1^I \theta_4^A + \left[\alpha_6^I + \alpha_8^I \rho_z\right] Z_y + \alpha_8^I \theta_4^Z + \alpha_3^I X_{1B}.$$

Now, I proceed to solve by introducing the market-clearing equation:

$$\frac{p_4}{\beta} = \left(\alpha_1^I \rho_A + \alpha_7^I\right) \tilde{A}_1 + \alpha_1^I \theta_4^A + \left[\alpha_6^I + \alpha_8^I \rho_z\right] Z_y + \alpha_8^I \theta_4^Z + \left[\alpha_3^I - \delta^{II} 2\right] X_{1B}.$$

$$\frac{p_4}{\beta} = \begin{bmatrix} \left(\alpha_1^I \rho_A + \alpha_7^I\right) \tilde{A}_1 + \alpha_1^I \theta_4^A + \left[\alpha_6^I + \alpha_8^I \rho_z\right] Z_y \\ +\alpha_8^I \theta_4^Z + \left[\alpha_3^I - \delta^{II} 2\right] \left(-\bar{p}_4^{-\rho} \tilde{Z}_4 + X_4 + \rho \bar{Z}_4' \bar{p}_4^{-\rho-1} p_t\right). \end{bmatrix}$$

$$\frac{p_4}{\beta} = \begin{bmatrix} \left(\alpha_1^I \rho_A + \alpha_7^I\right) \tilde{A}_1 + \alpha_1^I \theta_4^A + \left[\alpha_6^I + \alpha_8^I \rho_z\right] Z_y + \alpha_8^I \theta_4^Z \\ + \left[\alpha_3^I - \delta^{II} 2\right] \left(-\bar{p}_4^{-\rho} \left(Z_y + \varepsilon_4^z\right) + X_4 + \rho \bar{Z}_4' \bar{p}_4^{-\rho-1} p_t\right). \end{bmatrix}$$

$$\frac{p_4}{\beta} = \begin{bmatrix} \left(\alpha_1^I \rho_A + \alpha_7^I\right) \tilde{A}_1 + \left[\alpha_6^I + \alpha_8^I \rho_z\right] Z_y + \alpha_1^I \theta_4^A + \alpha_8^I \theta_4^Z \\ + \left[\alpha_3^I - \delta^{II} 2\right] \left(-\bar{p}_4^{-\rho} \left(Z_y + \varepsilon_4^z\right) + X_4 + \rho \bar{Z}_4' \bar{p}_4^{-\rho-1} p_t\right) \end{bmatrix}.$$

$$\frac{p_4}{\beta} = \begin{bmatrix} \left(\alpha_1^I \rho_A + \alpha_7^I\right) \tilde{A}_1 + \left[\alpha_6^I + \alpha_8^I \rho_z\right] Z_y + \alpha_1^I \theta_4^A + \alpha_8^I \theta_4^Z \\ - \left(\alpha_3^I - \delta^{II} 2\right) \bar{p}_4^{-\rho} Z_y - \left(\alpha_3 - \delta^{II} 2\right) \bar{p}_4^{-\rho} \varepsilon_4^z + \\ + \left[\alpha_3 - \delta^{II} 2\right] X_4 + \left[\alpha_3 - \delta^{II} 2\right] \rho \bar{Z}_4' \bar{p}_4^{-\rho-1} p_t \end{bmatrix}.$$

$$\frac{p_4}{\beta} - \left[\alpha_3^I - \delta^{II} 2\right] \rho \bar{Z}_4' \bar{p}_4^{-\rho-1} p_t = \begin{bmatrix} \left(\alpha_1^I \rho_A + \alpha_7^I\right) \tilde{A}_1 + \left[\alpha_6^I + \alpha_8^I \rho_z - \left(\alpha_3^I - \delta^{II} 2\right) \bar{p}_4^{-\rho}\right] Z_y \\ - \left(\alpha_3^I - \delta^{II} 2\right) \bar{p}_4^{-\rho} \varepsilon_4^z + \\ + \left[\alpha_3^I - \delta^{II} 2\right] X_4 + \alpha_1^I \theta_4^A + \alpha_8^I \theta_4^Z \end{bmatrix}.$$

$$\begin{aligned} p_4 \quad = \quad & \frac{\left(\alpha_1^I \rho_A + \alpha_7\right)}{\left[\frac{1}{\beta} - \left(\alpha_3^I - \delta^{II} 2\right) \rho \bar{Z}_4' \bar{p}_4^{-\rho-1}\right]} \tilde{A}_1 - \frac{\left(\alpha_3^I - \delta^{II} 2\right) \bar{p}_4^{-\rho}}{\left[\frac{1}{\beta} - \left(\alpha_3^I - \delta^{II} 2\right) \rho \bar{Z}_4' \bar{p}_4^{-\rho-1}\right]} \varepsilon_4^z \\ & + \frac{\left[\alpha_3^I - \delta^{II} 2\right]}{\left[\frac{1}{\beta} - \left(\alpha_3^I - \delta^{II} 2\right) \rho \bar{Z}_4' \bar{p}_4^{-\rho-1}\right]} X_4 + \frac{\alpha_1^I}{\left[\frac{1}{\beta} - \left(\alpha_3^I - \delta^{II} 2\right) \rho \bar{Z}_4' \bar{p}_4^{-\rho-1}\right]} \theta_4^A \\ & + \frac{\alpha_8^I}{\left[\frac{1}{\beta} - \left(\alpha_3^I - \delta^{II} 2\right) \rho \bar{Z}_4' \bar{p}_4^{-\rho-1}\right]} \theta_4^Z + \frac{\left[\alpha_6^I + \alpha_8^I \rho_z - \left(\alpha_3^I - \delta^{II} 2\right) \bar{p}_4^{-\rho}\right]}{\left[\frac{1}{\beta} - \left(\alpha_3^I - \delta^{II} 2\right) \rho \bar{Z}_4' \bar{p}_4^{-\rho-1}\right]} Z_y. \end{aligned}$$

$$\alpha_1^{IV} = \frac{\alpha_1^I \rho_A + \alpha_7^I}{\left[\frac{1}{\beta} - \left(\alpha_3^I - \delta^{II} 2\right) \rho \bar{Z}_4' \bar{p}_4^{-\rho-1}\right]}.$$

$$\alpha_2^{IV} = -\frac{\left(\alpha_3^I - \delta^{II} 2\right) \bar{p}_4^{-\rho}}{\left[\frac{1}{\beta} - \left(\alpha_3^I - \delta^{II} 2\right) \rho \bar{Z}_4' \bar{p}_4^{-\rho-1}\right]}.$$

$$\alpha_3^{IV} = \frac{\left[\alpha_3^I - \delta^{II}2\right]}{\left[\frac{1}{\beta} - \left(\alpha_3^I - \delta^{II}2\right)\rho \bar{Z}_4' \bar{p}_4^{-\rho-1}\right]}.$$

$$\alpha_4^{IV} = \frac{\alpha_1^I}{\left[\frac{1}{\beta} - \left(\alpha_3^I - \delta^{II}2\right)\rho \bar{Z}_4' \bar{p}_4^{-\rho-1}\right]}$$

$$\alpha_5^{IV} = \frac{\alpha_8^I}{\left[\frac{1}{\beta} - \left(\alpha_3^I - \delta^{II}2\right)\rho \bar{Z}_4' \bar{p}_4^{-\rho-1}\right]}$$

$$\alpha_8^{IV} = \frac{\left[\alpha_6^I + \alpha_8 \rho_z - \left(\alpha_3^I - \delta^{II}2\right)\bar{p}_4^{-\rho}\right]}{\left[\frac{1}{\beta} - \left(\alpha_3^I - \delta^{II}2\right)\rho \bar{Z}_4' \bar{p}_4^{-\rho-1}\right]}$$

17 Appendix Part 2

17.1 Data

	Futures		
Commodity	Ticker and Source	Length	Type
Soybean	S1 - Bloomberg	Jan 2002-Apr 2017	First generic future
Corn	C1 - Bloomberg	Jan 2002-Apr 2017	First generic future
Cotton	CT1 - Bloomberg	Jan 2002-Apr 2017	First generic future
Wheat	W1 - Bloomberg	Jan 2002-Apr 2017	First generic future
Gold	GC1 - Bloomberg	Jan 2002-Apr 2017	First generic future
Silver	SI1 - Bloomberg	Jan 2002-Apr 2017	First generic future
Oil	CL1 - Bloomberg	Jan 2002-Apr 2017	First generic future

	Harvest CDF		
Commodity	Source	Length	Description
Soybean	USDA	2011-2016	Fraction of harvest picked up per week
Corn	USDA	2011-2016	Fraction of harvest picked up per week
Cotton	USDA	2011-2016	Fraction of harvest picked up per week
Wheat	USDA	2011-2016	Fraction of harvest picked up per week
	Cash prices		
Corn - Illinois	USDA	1992-2016	Cash prices. Central Illinois
Corn - Iowa	USDA	1992-2016	Cash prices. South Central Iowa

17.2 Speculators Problem:

$$\underset{X_j, a_{j,t}, b(j,t)}{Max} \; E_j \left[\begin{array}{c} \left[\sum_{t=j}^{T} m_{j,t} S_t X_j (1-\delta)^{t-j} a_{j,t} + F(j,t) X_j (1-\delta)^{t-j} (1-a_{j,t}) + b(j,t) \right] \\ - \sum_{t=j}^{T} q(j,t) b(j,t) - S_j X_j \end{array} \right]$$

s.t. $\quad 0 \leq a_{j,t} \leq 1 \qquad X_j \geq 0.$

Notice that the problem is linear in its arguments, hence equilibrium results must keep this agent indifferent.

Set $\delta = 0$

$$\underset{X_{j,t},a_{j,t},b(j,t)}{Max} \; E_j \left[\begin{array}{c} \sum_{t=j}^{T} m_{j,t} \left(S_t X_{j,t} a_{j,t} + F(j,t) X_{j,t} (1 - a_{j,t}) + b(j,t) \right) \\ - \left(\sum_{t=j}^{T} q(j,t) b(j,t) + S_j X_{j,t} \right) \end{array} \right] + \sum_{t=j}^{T} \lambda_{j,t} X_{j,t}$$

s.t. $\quad 0 \leq a_{j,t} \leq 1$

FOC

$$E_j \left(m_{j,t} \left[S_t a_{j,t} + F(j,t)(1 - a_{j,t}) \right] \right) - S_j + \lambda_{j,t} = 0$$

$$E_j \left(m_{j,t} \left[S_t a_{j,t} + F(j,t)(1 - a_{j,t}) \right] \right) \leq S_j \text{ (with equality if } X_{j,t} > 0)$$

$$E_j \left[m_{j,t} \left(S_t X_{j,t} - F(j,t) X_{j,t} \right) \right] = 0$$

$$E_j(m_{j,t} S_t) = E_j(m_{j,t}) F(j,t) \qquad .$$

$$E_j \left(m_{j,t} \right) - q(j,t) = 0$$

$$Ej \left(m_{j,t} \right) = q(j,t)$$

$$\lambda_{j,t} \geq 0, X_{j,t} \geq 0$$

Combining

$$q(j,t) F(j,t) = E_j \left(m_{j,t} S_t \right) \leq S_j$$

Under an equivalent measure (as in Pirrong 2011):

$$q(j,t) F(j,t) = \hat{E}_j \left(S_t \right) \leq S_j$$

Scenario I future curves

No leftovers condition is now:

$$Q_0^s + y = \sum_{t=0}^{T} D_t(S_t) = \sum_{t=0}^{T} -dQ_t^S dt = \sum_{t=0}^{T} -dQ_0^S q^{\rho t} dt = -dQ_0^s \tfrac{1-q^{\rho(T+1)}}{1-q^{\rho}}$$

$$(Q_0^s + y)\,\frac{1-q^{\rho}}{1-q^{\rho(T+1)}} = -dQ_0^s \qquad \text{(A.1.a)}$$

Introducing A.1.a into 8.a and 8.b gives

$$S_0 = M(0,T)\hat{E}_0\left[(Q_0^s + y)^{-\rho}\right]$$

$$F(0,t) = M(t,T)\hat{E}_0\left[(Q_0^s + y)^{-\rho}\right] \qquad 0 \le t \le T$$

Scenario II future curves

$$S_0 = M(0,J){Q_0^s}^{-\rho}$$

$$F(0,t) = M(t,J){Q_0^s}^{-\rho}$$

$$F(J,J) = M(0,T-J)\hat{E}_0\left[y^{-\rho}\right]$$

$$F(J,t) = M(t-J,T-J)\hat{E}_0\left[y^{-\rho}\right] \qquad J \le t \le T$$

The previous equations shown in section II are just an extension of equations in section I. In order to obtain these, we can follow the same procedure shown above for section I. In this case for each interval, we can use expression A.1.a adapted to the harvest during that interval. Then we introduce into equation 8.a and 8.b to obtain final results.

17.3 Continuous case

The planner's problem is:

$$\underset{x}{Max}\ E_0\left[\int_0^T \exp(-rt)\,(U(x))\,dt\right]$$

st $\quad \dot{Q} = y - x \quad Q \geq 0 \quad Q(0) > 0 \quad$ where $y(t)$ indicates the harvest that is picked up in moment t,i.e., $y(t) = g(t)H$, where $g(.)$ is the pdf of the harvest cumulative function.

$Q(T) = 0$

Information will be revealed at time $0 < J \leq a$. Hence the problem can be divided in two:

$$\underset{x}{Max}\int_0^J \exp(-rt)\,(U(x))\,dt + \exp(-rJ)E_0\left[\underbrace{\int_J^T \exp(-r(t-J))\,(U(x))\,dt}_{V(Q(J))}\right]$$

st $\quad \dot{Q} = y - x \quad Q \geq 0 \quad Q(0) > 0$

$Q(T) = 0$

Step 1: Solve for $V(Q(J))$

$$V(Q(J)) = \underset{x}{Max}E_0\left[\int_J^T \exp(-r(t-J))\,(U(x))\,dt\right]$$

st $\quad \dot{Q} = y - x \quad Q \geq 0 \quad Q(J) > 0$

$Q(T) = 0$

Redefining $J = 0$ WLOG:

$$V(Q(J)) = \underset{x}{Max}\int_{0}^{T} \exp(-r(t))\,(U(x))\,dt$$

st $\quad \dot{Q} = y - x \qquad Q \geq 0 \qquad Q(0) > 0$

$Q(T) = 0$

Setting up the Hamiltonian's necessary and sufficient conditions:

$H = \exp(-rt)U(x) + \lambda(y - x) + \eta Q$

$H_x = \exp(-rt)U_x' - \lambda = 0 \qquad \Rightarrow \qquad \exp(-rt)Zx^{-\rho} = \lambda$

$\lambda' = -H_Q = \lambda\frac{r}{\rho} + \eta$

$\dot{Q} = y - x \qquad \eta Q = 0$

When the constraint is not active:

$$\frac{\dot{x}(t)}{x(t)} = -\frac{r}{\rho} \text{ for all } 0 \leq t \leq T \tag{A.1.b}$$

$\int_{0}^{t} \frac{\dot{x}}{x}dt = -\frac{r}{\rho}\int_{0}^{t} dt$

$\log x(t) - \log(x(0)) = -\frac{r}{\rho}t$

$\log x(t) = \log(x(0)) - \frac{r}{\rho}t$

$x(t) = \exp\left[\log(x(0)) - \frac{r}{\rho}t\right]$

$$x(t) = x(0)\exp(-\frac{r}{\rho}t) \tag{A.2}$$

If the constraint never becomes active, we would get:

$\dot{Q} = y - x$

$$\int_0^T \dot{Q}dt = \int_0^T (y-x)dt$$

$$\underbrace{Q(T)} - Q(0) = G(T)H - \int_0^T xdt$$

$$x(0)\int_0^T \exp(-\tfrac{r}{\rho}t)dt = H + Q(0)$$

$$x(0)\tfrac{\rho}{r}\left[1-\exp(-\tfrac{r}{\rho}T)\right] = H + Q(0)$$

$$\Rightarrow x(0) = \frac{r}{\rho}\frac{H+Q(0)}{\left[1-\exp(-\frac{r}{\rho}T)\right]} \tag{A.3}$$

$$\dot{Q}(t) = y(t) - \frac{r}{\rho}\frac{H+Q(0)}{\left[1-\exp(-\frac{r}{\rho}T)\right]}\exp(-\frac{r}{\rho}t) \tag{A.4}$$

$$Q(t) = Q(0) + \int_0^t \dot{Q}(t)dt$$

$$Q(t) - Q(0) = \int_0^t \left[y(t) - \tfrac{r}{\rho}\tfrac{H+Q(0)}{\left[1-\exp(-\frac{r}{\rho}T)\right]}\exp(-\tfrac{r}{\rho}t)\right]dt$$

$$Q(t) - Q(0) = \int_0^t y(t)dt - \tfrac{r}{\rho}\tfrac{H+Q(0)}{\left[1-\exp(-\frac{r}{\rho}T)\right]}\int_0^t \exp(-\tfrac{r}{\rho}t)dt$$

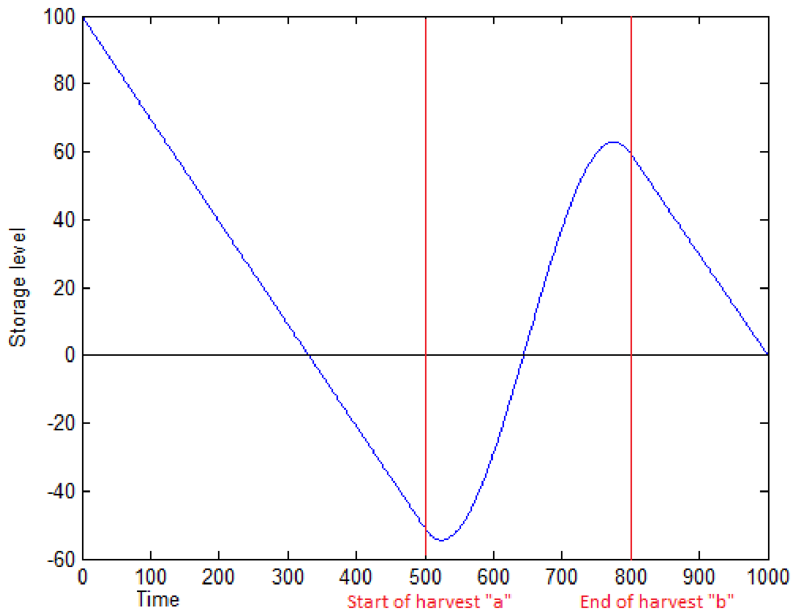
$$Q(t) - Q(0) = G(t)H - \tfrac{r}{\rho}\tfrac{H+Q(0)}{\left[1-\exp(-\frac{r}{\rho}T)\right]}\tfrac{\rho}{r}\left[1-\exp(-\tfrac{r}{\rho}t)\right]$$

$$Q(t) - Q(0) = G(t)H - \tfrac{H+Q(0)}{\left[1-\exp(-\frac{r}{\rho}T)\right]}\left[1-\exp(-\tfrac{r}{\rho}t)\right]$$

$$Q(t) = G(t)H + Q(0) - \frac{H + Q(0)}{\left[1 - \exp(-\frac{r}{\rho}T)\right]}\left[1 - \exp(-\frac{r}{\rho}t)\right] \quad \text{for } 0 \leq t \leq T$$

For H big enough, Q will reach the zero bound at some moment A and will then return to positive values for some moment B :

Figure A.I



Clearly, the solution in Figure A.I is not feasible, hence for H big enough we must find the optimal breaking points A, B.

If the constraint is active

$\lambda'(t) = \lambda(t)\frac{r}{\rho} + \eta(t)$

$x(t) = y(t) \qquad \Rightarrow \qquad \lambda(t) = \exp(-rt)Zy(t)^{-\rho}$

Let us define A as the first moment when the restriction become active, hence $Q(A) = 0$:

$\dot{Q} = y - x$

$$\int_0^A \dot{Q}dt = \int_0^A (y - x)dt$$

$$\underbrace{Q(A)}_{=0} - Q(0) = G(A)H - \int_0^A x dt$$

$$x(0)\int_0^A \exp(-\tfrac{r}{\rho}t)dt = G(A)H + Q(0)$$

$$x(0)\tfrac{\rho}{r}\left[1 - \exp(-\tfrac{r}{\rho}A)\right] = G(A)H + Q(0)$$

$$\Rightarrow x(0) = \frac{r}{\rho}\frac{G(A)H + Q(0)}{\left[1 - \exp(-\frac{r}{\rho}A)\right]} \tag{A.3.A}$$

$$\dot{Q}(t) = y(t) - \frac{r}{\rho}\frac{G(A)H + Q(0)}{\left[1 - \exp(-\frac{r}{\rho}t)\right]}\exp(-\frac{r}{\rho}t) \tag{A.4.A}$$

For $A < t < B$

$Q(t) = 0, \qquad \dot{Q}(t) = 0$

For $B < t \leq T$, we have

$$\frac{\dot{x}(t)}{x(t)} = -\frac{r}{\rho} \text{ for all } B \leq t \leq T$$

$$\int_B^t \tfrac{\dot{x}}{x}dt = -\tfrac{r}{\rho}\int_B^t dt$$

$$\log x(t) - \log(x(B)) = -\tfrac{r}{\rho}(t - B)$$

$$\log x(t) = \log(x(B)) - \tfrac{r}{\rho}(t - B)$$

$$x(t) = \exp\left[\log(x(B)) - \tfrac{r}{\rho}(t - B)\right]$$

$$x(t) = x(B)\exp(-\frac{r}{\rho}(t-B))$$

$$\dot{Q} = y - x$$

$$\int\limits_B^T \dot{Q}dt = \int\limits_B^T (y-x)dt$$

$$\underbrace{Q(T)}_{=0} - \underbrace{Q(B)}_{=0} = (1-G(B))H - \int\limits_B^T xdt$$

$$x(B)\int\limits_B^T \exp(-\tfrac{r}{\rho}(t-B))dt = (1-G(B))H$$

$$x(0)\tfrac{\rho}{r}\left[1-\exp(-\tfrac{r}{\rho}(T-B)\right] = (1-G(B))H$$

$$x(B) = \frac{r}{\rho}\frac{(1-G(B))H}{\left[1-\exp(-\frac{r}{\rho}(T-B)\right]} \tag{A.3.B}$$

$$\dot{Q}(t) = y(t) - \frac{r}{\rho}\frac{(1-G(B))H}{1-\exp(-\frac{r}{\rho}(T-B))}\exp(-\frac{r}{\rho}(t-B)) \tag{A.4.B}$$

$$Q(t) = Q(B) + \int\limits_B^t \dot{Q}(t)dt$$

$$Q(t) = \int\limits_B^t \dot{Q}(t)dt$$

$$Q(t) = \int\limits_B^t \left[y(t) - \tfrac{r}{\rho}\tfrac{(1-G(B))H}{\left[1-\exp(-\frac{r}{\rho}(T-B))\right]}\exp(-\tfrac{r}{\rho}(t-B))\right]dt$$

$$Q(t) = \int\limits_B^t y(t)dt - \tfrac{r}{\rho}\tfrac{(1-G(B))H}{\left[1-\exp(-\frac{r}{\rho}(T-B))\right]}\int\limits_B^t \exp(-\tfrac{r}{\rho}(t-B))dt$$

$$Q(t) = (G(t) - G(B))H - \frac{(1-G(B))H}{\left[1-\exp(-\frac{r}{\rho}(T-B))\right]}\left[1 - \exp(-\frac{r}{\rho}(t - B)\right]$$

Therefore, we have:

$$Q(t)\left\{\begin{array}{ll} Q(t) = G(t)H + Q(0) - [G(A)H + Q(0)]\,\frac{\left[1-\exp(-\frac{r}{\rho}t)\right]}{\left[1-\exp(-\frac{r}{\rho}A)\right]} & \text{for } 0 \leq t \leq A \\ Q(t) = 0 & \text{for } A \leq t \leq B \\ Q(t) = (1 - G(t))H - [(1 - G(B))\,H]\,\frac{\left[1-\exp(-\frac{r}{\rho}(T-t))\right]}{\left[1-\exp(-\frac{r}{\rho}(T-B))\right]} & \text{for} B \leq t \leq T \end{array}\right\} \quad \text{A.5}$$

Since λ is continuous, then x is continuous also. Therefore $\dot{Q}$ is tangent in the constraint moments A and B. We must have

$$\dot{Q}(A^-) = y(A^+) - x(A^+) \qquad \Rightarrow \qquad 0 = y(A^+) - x(A^+) \qquad \Rightarrow \qquad y(A^-) = x(A^+)$$

$$\Rightarrow \qquad g(A^*)H = \frac{r}{\rho}\frac{G(A)H + Q(0)}{\left[1 - \exp(-\frac{r}{\rho}A^*)\right]}\exp(-\frac{r}{\rho}A^*) \text{ determines implicitly the value of } A^* \tag{A.6}$$

For B,

$$0 = \dot{Q}(B^+) = y(B^-) - x(B^-) \qquad \Rightarrow \qquad y(B^-) = x(B^-)$$

$$\Rightarrow \qquad g(B^*) = \frac{r}{\rho}\frac{(1 - G(B^*))}{\left[1 - \exp(-\frac{r}{\rho}(T - B^*))\right]} \qquad \text{determines implicitly the value of } B^* \tag{A.7}$$

Notice that the value of B^* is independent of the size of H.

Therefore, if $A^* < B^*$ then the market will be in backwardation. Otherwise, solution will be $A^* = B^* = T$, that is, the market will be in contango, since the non-negative constraint on storage will never be active.

For simplicity, assume $r = 0$

$x(t) = x(0) = x(A) = \frac{G(A)H+Q(0)}{A}$

$$\begin{aligned} Q(t) &= G(t)H + Q(0) - \frac{[G(A)H + Q(0)]}{A}t && \text{for } 0 \le t \le A \\ Q(t) &= 0 && \text{for } A \le t \le B \\ Q(t) &= (1 - G(t))H - \frac{[(1 - G(B))\,H]}{T - B}(T - t) && \text{for} B \le t \le T \end{aligned}$$

Solution will be

$\frac{G(A)H+Q(0)}{A} = g(A)H \qquad \Rightarrow \qquad g(A)A - G(A) = \frac{Q(0)}{H}$

$g(B) = \frac{[1-G(B)]}{T-B}$

For H small enough, equation A.6 might never be satisfied. In that case, $A^* = B^* = T$, that is, future prices will be in contango.

Therefore, I can map A^*, B^* as function of H

$A^* = A(H) \qquad B^* = B(H)$

The final value function is then

If backwardation:

$$V(Q(J)) = \underbrace{\int_J^{A^*} \exp(-rt)Z\frac{x(t)^{1-\rho}}{1-\rho}dt}_{V^1} + \underbrace{\int_{A^*}^{B^*} \exp(-rt)Z\frac{x(t)^{1-\rho}}{1-\rho}dt}_{V^2} + \underbrace{\int_{B^*}^{T} \exp(-rt)Z\frac{x(t)^{1-\rho}}{1-\rho}dt}_{V^3}$$

$$V^1 = z\frac{\left[\frac{r}{\rho}\frac{G(A^*)H+Q(0)}{\left[1-\exp(-\frac{r}{\rho}(A^*-J))\right]}\right]^{1-\rho}}{1-\rho}\int_J^{A^*} \exp(-r(t-J) - \tfrac{r}{\rho}(t-J)(1-\rho))dt$$

$$V^1 = z\frac{\left[\frac{r}{\rho}\frac{G(A^*)H+Q(0)}{\left[1-\exp(-\frac{r}{\rho}(A^*-J))\right]}\right]^{1-\rho}}{1-\rho}\int_J^{A^*}\exp(-r(t-J)-\frac{r}{\rho}(t-J)(1-\rho))dt$$

$$V^1 = z\frac{\left[\frac{r}{\rho}\frac{G(A^*)H+Q(0)}{\left[1-\exp(-\frac{r}{\rho}(A^*-J))\right]}\right]^{1-\rho}}{1-\rho}\int_J^{A^*}\exp(-rt+rJ-\frac{r}{\rho}t+\frac{r}{\rho}J+rt-rJ))dt$$

$$V^1 = z\frac{\left[\frac{r}{\rho}\frac{G(A^*)H+Q(0)}{\left[1-\exp(-\frac{r}{\rho}(A^*-J))\right]}\right]^{1-\rho}}{1-\rho}\exp(\frac{r}{\rho}J)\int_J^{A^*}\exp(-\frac{r}{\rho}t))dt$$

$$V^1 = z\frac{\left[\frac{r}{\rho}\frac{G(A^*)H+Q(0)}{\left[1-\exp(-\frac{r}{\rho}(A^*-J))\right]}\right]^{1-\rho}}{1-\rho}\exp(\frac{r}{\rho}J)\int_J^{A^*}\exp(-\frac{r}{\rho}t))dt$$

$$V^1 = z\frac{\left[\frac{r}{\rho}\frac{G(A^*)H+Q(0)}{\left[1-\exp(-\frac{r}{\rho}(A^*-J))\right]}\right]^{1-\rho}}{1-\rho}\frac{\rho}{r}\exp(\frac{r}{\rho}J)\left[-\exp(-\frac{r}{\rho}A^*)+\exp(-\frac{r}{\rho}J)\right]$$

$$V^1(Q(J),H) = \frac{z}{1-\rho}\frac{\rho}{r}\left[\frac{r}{\rho}\frac{G(A^*)H+Q(J)}{\left[1-\exp(-\frac{r}{\rho}(A^*-J))\right]}\right]^{1-\rho}\left[1-\exp(-\frac{r}{\rho}(A^*-J))\right] \qquad A^* = A^*(Q(J))$$

$$V^2(H,\bar{Q}) = \int_{A^*}^{B^*}\exp(-rt)Z\frac{x(t)^{1-\rho}}{1-\rho}dt$$

$$V^2(H,\bar{Q}) = \int_{A^*}^{B^*}\exp(-rt)Z\frac{[g(t)H]^{1-\rho}}{1-\rho}dt$$

$$V^2(H,\bar{Q}) = \exp(-rA^*)\frac{z}{1-\rho}H^{1-\rho}\int_{A^*}^{B^*}\exp(-r(t-A^*))g(t)^{1-\rho}dt$$

$$V^3(H) = \exp(-rB^*)\frac{z}{1-\rho}\frac{\rho}{r}\left[\frac{r}{\rho}\frac{1-G(B^*)H}{\left[1-\exp(-\frac{r}{\rho}(T-B^*))\right]}\right]^{1-\rho}\left[1-\exp(-\frac{r}{\rho}(T-B^*))\right]$$

If contango:

$$V(Q(J),H) = \frac{z}{1-\rho}\frac{\rho}{r}\left[\frac{r}{\rho}\frac{H+Q(J)}{\left[1-\exp(-\frac{r}{\rho}(T-J))\right]}\right]^{1-\rho}\left[1-\exp(-\frac{r}{\rho}(T-J))\right]$$

Now we can return back to our initial problem:

$$\underset{x}{Max}\underbrace{\int_0^J \exp(-rt)\,(U(x))\,dt}_{N(Q(0))} + \exp(-rJ)E_0\left[\underbrace{\int_J^T \exp(-r(t-J))\,(U(x))\,dt}_{V(Q(J))}\right]$$

st $\quad \dot{Q} = y - x \quad Q \geq 0 \quad Q(0) > 0$

$Q(T) = 0$

We can solve for $N(Q(0))$ conditional on some terminal value $Q(J) = \bar{Q}$

$$N(Q(0)/\bar{Q}) = \underset{x}{Max}\int_0^J \exp(-rt)\,(U(x))\,dt$$

st $\quad \dot{Q} = y - x \quad Q \geq 0 \quad Q(0) > 0$

$Q(J) = \bar{Q}$

Since $y(t) = 0$ for this interval, the restriction will never be active. Hence we can solve for a standard hamiltonian as above, obtaining identical results as in A.2, A.3, A.4:

$$x(t) = x(0)\exp(-\frac{r}{\rho}t)$$

$\dot{Q} = -x$

$$\int_0^J \dot{Q}dt = \int_0^J -xdt$$

$$\underbrace{Q(J)}_{=\bar{Q}} - Q(0) = -\int_0^J x dt$$

$$x(0)\int_0^J \exp(-\tfrac{r}{\rho}t)dt = Q(0) - \bar{Q}$$

$$x(0)\tfrac{\rho}{r}\left[1 - \exp(-\tfrac{r}{\rho}J)\right] = Q(0) - \bar{Q}$$

$$x(0) = \frac{r}{\rho}\frac{Q(0) - \bar{Q}}{\left[1 - \exp(-\tfrac{r}{\rho}J)\right]} \tag{B.3}$$

$$\dot{Q}(t) = -\frac{r}{\rho}\frac{Q(0) - \bar{Q}}{\left[1 - \exp(-\tfrac{r}{\rho}J)\right]}\exp(-\frac{r}{\rho}t) \tag{B.4}$$

$$N(Q(0)) = \int_0^J \exp(-rt)\left(U(x)\right)dt$$

$$= \int_0^J \exp(-rt) Z \frac{\left(\frac{r}{\rho}\frac{Q(0)-\bar{Q}}{\left[1-\exp(-\frac{r}{\rho}J)\right]}\right)^{1-\rho}}{1-\rho}\exp(-\tfrac{r}{\rho}t(1-\rho))dt$$

$$= Z\frac{\left(\frac{r}{\rho}\frac{Q(0)-\bar{Q}}{\left[1-\exp(-\frac{r}{\rho}J)\right]}\right)^{1-\rho}}{1-\rho}\int_0^J \exp(-rt - \tfrac{r}{\rho}t + rt)dt$$

$$= Z\frac{\left(\frac{r}{\rho}\frac{Q(0)-\bar{Q}}{\left[1-\exp(-\frac{r}{\rho}J)\right]}\right)^{1-\rho}}{1-\rho}\int_0^J \exp(-\tfrac{r}{\rho}t)dt$$

$$N(Q(0)) = \frac{Z}{1-\rho}\left(\frac{r}{\rho}\frac{Q(0) - \bar{Q}}{\left[1 - \exp(-\frac{r}{\rho}J)\right]}\right)^{1-\rho}\frac{\rho}{r}\left[1 - \exp(-\frac{r}{\rho}J)\right]$$

The final solution will be given by:

$$\max_{\bar{Q}} N(Q(0)/\bar{Q}) + \exp(-rJ)E_0\left[V(\bar{Q})\right]$$

st $\qquad \bar{Q} \leq Q(0)$

FOC

$-\frac{\partial N(.)}{\partial \bar{Q}} = \exp(-rJ)E_0\left[\frac{\partial V(\bar{Q})}{\partial \bar{Q}}\right]$

For values of $\bar{Q}$ that give backwardation:

$\frac{\partial V(\bar{Q})}{\partial \bar{Q}} = \frac{\partial V^1(\bar{Q})}{\partial \bar{Q}} + \frac{\partial V^1(\bar{Q})}{\partial A^*}\frac{\partial A^*}{\partial \bar{Q}} + \frac{\partial V^2(\bar{Q})}{\partial A^*}\frac{\partial A^*}{\partial \bar{Q}}$

$\frac{\partial V(\bar{Q})}{\partial \bar{Q}} = \frac{\partial V^1(\bar{Q})}{\partial \bar{Q}} + (\frac{\partial V^1(\bar{Q})}{\partial A^*} + \frac{\partial V^2(\bar{Q})}{\partial A^*})\frac{\partial A^*}{\partial \bar{Q}}$

Recall that

$\frac{\partial V^1(\bar{Q})}{\partial A^*} + \frac{\partial V^2(\bar{Q})}{\partial A^*} = \exp(-rA^*)Z\frac{1}{1-\rho}x(A^{*-})^{1-\rho} - \exp(-rA^*)Z\frac{1}{1-\rho}x(A^{*+})^{1-\rho}$

Since optimal x is continuous in A^*, we have $\frac{\partial V^1(\bar{Q})}{\partial A^*} + \frac{\partial V^2(\bar{Q})}{\partial A^*} = 0$

$$\frac{\partial V(\bar{Q})}{\partial \bar{Q}} = \frac{\partial V^1(\bar{Q})}{\partial \bar{Q}}$$

For values of $\bar{Q}$ that give backwardation:

$\frac{\partial V(\bar{Q})}{\partial \bar{Q}} = Z\frac{\rho}{r}\left[\frac{r}{\rho}\frac{H+Q(J)}{\left[1-\exp(-\frac{r}{\rho}(T-J))\right]}\right]^{-\rho} \left[1-\exp(-\frac{r}{\rho}(T-J))\right]\frac{r}{\rho}\frac{1}{\left[1-\exp(-\frac{r}{\rho}(T-J))\right]}$

$$\frac{\partial V(\bar{Q})}{\partial \bar{Q}} = Z\left[\frac{r}{\rho}\frac{H+Q(J)}{\left[1-\exp(-\frac{r}{\rho}(T-J))\right]}\right]^{-\rho}$$

Therefore
$1 = E_0\left[\frac{\partial V(\bar{Q})}{\partial \bar{Q}}\frac{\exp(-rJ)}{Z\left(\frac{\rho}{r}\left[1-\exp(-\frac{r}{\rho}J)\right]\right)^{\rho}\left(Q(0)-\bar{Q}\ \right)^{-\rho}}\right]$

Backwardation points:

$\frac{\partial V^1(\bar{Q})}{\partial \bar{Q}} = Z\left(\frac{\rho}{r}\left[1-\exp(-\frac{r}{\rho}(A^*-J))\right]\right)^{\rho}\left[G(A^*)H+\bar{Q}\right]^{-\rho}$

$$\frac{\partial V^1(\bar{Q})}{\partial \bar{Q}} \frac{\exp(-rJ)}{Z\left(\frac{\rho}{r}\left[1-\exp(-\frac{r}{\rho}J)\right]\right)^{\rho}\left(Q(0)-\bar{Q}\ \right)^{-\rho}} = Z\frac{\left[\left(\frac{\rho}{r}\left[1-\exp(-\frac{r}{\rho}(A^*-J))\right]\right)^{\rho}\left[G(A^*)H+\bar{Q}\right]^{-\rho}\ \right]\exp(-rJ)}{Z\left(\frac{\rho}{r}\left[1-\exp(-\frac{r}{\rho}J)\right]\right)^{\rho}\left(Q(0)-\bar{Q}\ \right)^{-\rho}}$$

$$= \left[\left(\left[1-\exp(-\frac{r}{\rho}(A^*-J))\right]\right)^{\rho}\left[G(A^*)H+\bar{Q}\right]^{-\rho}\ \right]\frac{\exp(-rJ)}{\left(\left[1-\exp(-\frac{r}{\rho}J)\right]\right)^{\rho}\left(Q(0)-\bar{Q}\ \right)^{-\rho}}$$

$$1 = \exp(-rJ)\left[\left(\frac{1-\exp(-\frac{r}{\rho}(A^*-J))}{1-\exp(-\frac{r}{\rho}J)}\frac{Q(0)-\bar{Q}}{G(A^*)H+\bar{Q}(J)}\right)^{\rho}\right] \tag{B.6.A}$$

For contango points:

$$\frac{\partial V(\bar{Q})}{\partial \bar{Q}} \frac{\exp(-rJ)}{Z\left(\frac{\rho}{r}\left[1-\exp(-\frac{r}{\rho}J)\right]\right)^{\rho}\left(Q(0)-\bar{Q}\ \right)^{-\rho}} = Z\left[\frac{r}{\rho}\frac{H+Q(J)}{\left[1-\exp(-\frac{r}{\rho}(T-J))\right]}\right]^{-\rho}\frac{\exp(-rJ)}{Z\left(\frac{\rho}{r}\left[1-\exp(-\frac{r}{\rho}J)\right]\right)^{\rho}\left(Q(0)-\bar{Q}\ \right)^{-\rho}}$$

$$\frac{\partial V(\bar{Q})}{\partial \bar{Q}} \frac{\exp(-rJ)}{Z\left(\frac{\rho}{r}\left[1-\exp(-\frac{r}{\rho}J)\right]\right)^{\rho}\left(Q(0)-\bar{Q}\ \right)^{-\rho}} = \left[\frac{H+Q(J)}{\left[1-\exp(-\frac{r}{\rho}(T-J))\right]}\right]^{-\rho}\frac{\exp(-rJ)}{\left(\left[1-\exp(-\frac{r}{\rho}J)\right]\right)^{\rho}\left(Q(0)-\bar{Q}\ \right)^{-\rho}}$$

$$1 = \left[\frac{\left[1-\exp(-\frac{r}{\rho}(T-J))\right]}{\left[1-\exp(-\frac{r}{\rho}J)\right]}\frac{Q(0)-\bar{Q}}{H+Q(J)}\right]^{\rho}\exp(-rJ) \tag{B.6.B}$$

So, finally:

$$1 = \exp(-rJ)E_0\left[\begin{array}{c} \left(\frac{\left[1-\exp(-\frac{r}{\rho}(T-J))\right]}{\left[1-\exp(-\frac{r}{\rho}J)\right]}\frac{Q(0)-\bar{Q}}{H+Q(J)}\right)^{\rho} I_{\{A^*=T\}} \\ +\left(\frac{1-\exp(-\frac{r}{\rho}(A^*-J))}{1-\exp(-\frac{r}{\rho}J)}\frac{Q(0)-\bar{Q}}{G(A^*)H+\bar{Q}(J)}\right)^{\rho} I_{\{A^*<B^*\}} \end{array}\right] \tag{B.6.C}$$

... determines implicitly the optimal amount $\bar{Q}^* = Q(J)$.

Hence, given expectations, the path is determined for $0 \leq t < J$ by

$$x(t) = \frac{r}{\rho}\frac{Q(0)-\bar{Q}^*}{\left[1-\exp(-\frac{r}{\rho}J)\right]}\exp(-\frac{r}{\rho}t) \tag{B.3.A}$$

$$\dot{Q}(t) = -\frac{r}{\rho}\frac{Q(0) - \bar{Q}^*}{\left[1 - \exp(-\frac{r}{\rho}J)\right]}\exp(-\frac{r}{\rho}t) \tag{B.4.A}$$

and equation B.6 determines $\bar{Q}^*$.

Once $t = J$ is reached, information is fully revealed and H is no longer random.

Given this, consumption and storage path is given by:

for $0 < t < J$

$$x(t) = \frac{r}{\rho}\frac{Q(0) - \bar{Q}^*}{\left[1 - \exp(-\frac{r}{\rho}J)\right]}\exp(-\frac{r}{\rho}t) \tag{B.3.A}$$

$$\dot{Q}(t) = -\frac{r}{\rho}\frac{Q(0) - \bar{Q}^*}{\left[1 - \exp(-\frac{r}{\rho}J)\right]}\exp(-\frac{r}{\rho}t) \tag{B.4.A}$$

for $J \le t \le A^*$

$$x(t) = \frac{r}{\rho}\frac{\bar{Q}^* + G(A^*)H}{\left[1 - \exp(-\frac{r}{\rho}(A^* - J))\right]}\exp(-\frac{r}{\rho}(t - J)) \tag{B.3.B}$$

$$\dot{Q}(t) = g(t)H - \frac{r}{\rho}\frac{\bar{Q}^* + G(A^*)H}{\left[1 - \exp(-\frac{r}{\rho}(A^* - J))\right]}\exp(-\frac{r}{\rho}(t - J)) \tag{B.4.B}$$

for $A^* \le t \le B^*$

$$x(t) = g(t)H \tag{B.3.C}$$

$$\dot{Q}(t) = Q(t) = 0 \tag{B.4.C}$$

for $B^* \le t \le T$

$$x(t) = \frac{r}{\rho} \frac{(1 - G(B^*))H}{\left[1 - \exp(-\frac{r}{\rho}(T - B^*))\right]} \exp(-\frac{r}{\rho}(T - B^*)) \tag{B.3.D}$$

$$\dot{Q}(t) = g(t)H - \frac{r}{\rho} \frac{(1 - G(B^*))H}{\left[1 - \exp(-\frac{r}{\rho}(T - B^*))\right]} \exp(-\frac{r}{\rho}(T - B^*)) \tag{B.4.D}$$

17.3.1 Simple example

Assume $r = 0$. This will make algebra much easier without losing the main properties of the model.

Using L'Hopital, we get:

If backwardation

$$x(t) = \begin{array}{ll} \frac{Q(0)-\bar{Q}(J)}{J} & 0 \leq t \leq J \\ \frac{G(A^*)H+\bar{Q}(J)}{A^*-J} & J \leq t \leq A^* \\ g(t)H & A^* \leq t \leq B^* \\ \frac{(1-G(B^*))H}{T-B^*} & B^* \leq t \leq T \end{array}$$

If contango

$$x(t) = \begin{array}{ll} \frac{Q(0)-\bar{Q}(J)}{J} & 0 \leq t \leq J \\ \frac{H+\bar{Q}(J)}{T-J} & J \leq t \leq T \end{array}$$

A^* is determined by:

$G(A^*) + \frac{\bar{Q}(J)}{H} = g(A^*)\,(A^* - J)$ if backwardation

$A^* = T \qquad$ if contango

Using implicit function theorem:

$$\frac{\partial A^*}{\partial \bar{Q}} = -\frac{\frac{1}{H}}{g(A^*) - \frac{\partial g(A)}{\partial A}(A-J) - g(A)} = \frac{\frac{1}{H}}{\frac{\partial g(A)}{\partial A}(A-J)} = \frac{1}{H} \frac{1}{\frac{\partial g(A)}{\partial A}(A-J)} > 0$$

Hence:

$$V^1(.) = \int_J^{A^*} z \frac{\left(\frac{G(A^*)H+\bar{Q}(J)}{A^*-J}\right)^{1-\rho}}{1-\rho} dt$$

$V^1(.) = \frac{z}{1-\rho}\left(\frac{G(A^*)H+\bar{Q}(J)}{A^*-J}\right)^{1-\rho}\int\limits_J^{A^*} dt$

$V^1(.) = \frac{z}{1-\rho}\left(\frac{G(A^*)H+\bar{Q}(J)}{A^*-J}\right)^{1-\rho}(A^*-J)$

$V^1(.) = z\frac{(A^*-J)^\rho}{1-\rho}\left(G(A^*)H+\bar{Q}(J)\right)^{1-\rho}$

$$V^1(.) = z\frac{(A^*-J)^\rho}{1-\rho}\left(G(A^*)H+\bar{Q}(J)\right)^{1-\rho}$$

$V^2(.) = \int\limits_{A^*}^{B^*} Z\frac{[g(t)H]^{1-\rho}}{1-\rho}dt$

$$V^2(.) = Z\frac{H^{1-\rho}}{1-\rho}\int\limits_{A^*}^{B^*} g(t)^{1-\rho}dt$$

If contango

$V(.) = \int\limits_J^T Z\frac{\left(\frac{H+\bar{Q}(J)}{T-J}\right)^{1-\rho}}{1-\rho}dt$

$V(.) = Z\frac{\left(\frac{H+\bar{Q}(J)}{T-J}\right)^{1-\rho}}{1-\rho}\int\limits_J^T dt$

$V(.) = Z\frac{\left(\frac{H+\bar{Q}(J)}{T-J}\right)^{1-\rho}}{1-\rho}(T-J)$

$$V(H,\bar{Q}(J)) = \frac{Z}{1-\rho}(T-J)^\rho\left(H+\bar{Q}(J)\right)^{1-\rho}$$

$$N(Q(0)) = \frac{Z}{1-\rho}\left[Q(0)-\bar{Q}(J)\right]^{1-\rho}J^\rho$$

$-\frac{\partial N(.)}{\partial \bar{Q}} = E_0\left[\frac{\partial V^1(\bar{Q})}{\partial \bar{Q}}I_{\{A^*<B^*\}} + \frac{\partial V(\bar{Q})}{\partial \bar{Q}}I_{\{A^*=T\}}\right]$

$$ZJ^{\rho}\left[Q(0)-\bar{Q}(J)\right]^{-\rho}=E_0\left[\begin{array}{c} Z\left(A^{*}-J\right)^{\rho}\left(G(A^{*})H+\bar{Q}(J)\right)^{-\rho}I_{\{A^{*}<B^{*}\}} \\ +Z(T-J)^{\rho}\left(H+\bar{Q}(J)\right)^{-\rho}I_{\{A^{*}=T\}} \end{array}\right]$$

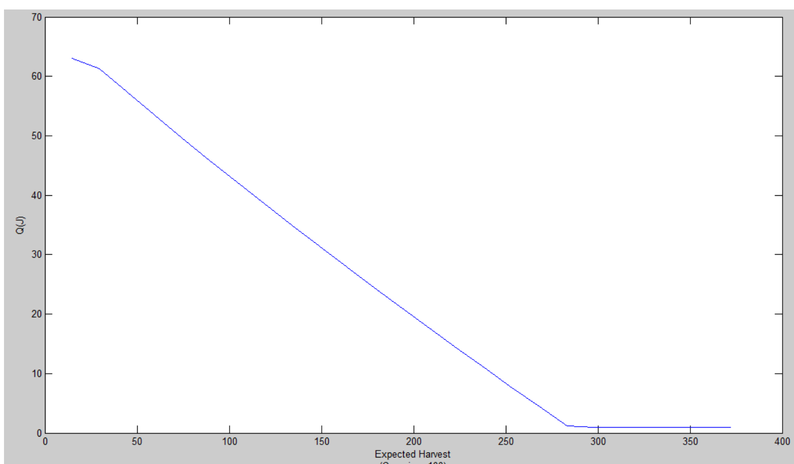
$$J^{\rho}\left[Q(0)-\bar{Q}(J)\right]^{-\rho}=E_0\left[\left(\tfrac{A^{*}-J}{G(A^{*})H+\bar{Q}(J)}\right)^{\rho}I_{\{A^{*}<B^{*}\}}+\left(\tfrac{T-J}{H+\bar{Q}(J)}\right)^{\rho}I_{\{A^{*}=T\}}\right]$$

$$1=E_0\left[\left(\tfrac{A^{*}-J}{J}\tfrac{Q(0)-\bar{Q}(J)}{G(A^{*})H+\bar{Q}(J)}\right)^{\rho}I_{\{A^{*}<B^{*}\}}+\left(\tfrac{T-J}{J}\tfrac{Q(0)-\bar{Q}(J)}{H+\bar{Q}(J)}\right)^{\rho}I_{\{A^{*}=T\}}\right]$$

$$1=E_0\left[\left(\tfrac{1}{g(A)J}(\tfrac{Q(0)-\bar{Q}}{H})\right)^{\rho}I_{\{A^{*}<B^{*}\}}+\left(\tfrac{T-J}{J}\tfrac{Q(0)-\bar{Q}(J)}{H+\bar{Q}(J)}\right)^{\rho}I_{\{A^{*}=T\}}\right]$$

$$1=E_0\left[\left(\frac{1}{g(A)J}(\frac{Q(0)-\bar{Q}}{H})\right)^{\rho}I_{\{A^{*}<B^{*}\}}+\left(\frac{T-J}{J}\frac{Q(0)-\bar{Q}(J)}{H+\bar{Q}(J)}\right)^{\rho}I_{\{A^{*}=T\}}\right] \tag{B.5}$$

Figure A.II



Simulation for $r = 0$

$\bar{Q}$ is continuous and changes when expectations change

Hence, when expectations related to the harvest increase(decrease), $\bar{Q}$ decreases, increasing $x(t)$ for $0 < t < J$, at the same time $E(x(t))$ increases for $B \leq t < T$ (post-harvest consumption). If prices are the inverse of demand, then we have a positive high correlation induced through supply-side information shocks.

17.4 Decentralization

Household's problem

$$\underset{x}{Max}\ E_0\left[\int_0^T \exp(-rt)\left(U(c)\right)+m(t)dt\right]$$

st $\qquad Sc+m=Y$

c : commodity good consumption

m : numeraire good

Y : income flow (assumed constant)

Since there are no state variables, solution is pointwise:

$$L=E_0\left[\int_0^T \left[\exp(-rt)\left(U(c)\right)+m(t)+\lambda(t)(Y-m(t)-S(t)c(t))\right]dt\right]$$

FOC

$\exp(-rt)Zc^{-\rho}=\lambda S$

$\lambda=1$

$m=Y-Sc$

$$Zc(t)^{-\rho}=S(t) \tag{C.1}$$

Combining the usual $Q\geq 0$ non-negative contraint, equations C.1 and C.2 we can replicate the planner's solution into a market equilibrium with prices $S(t)=\lambda(t), c(t)=x(t)$

A representative speculator interacts in the market. This embodies all agents in the economy: both farmers and financial agents, any of them can speculate. They do so by buying, selling and storing physical commodities and signing future contracts $F(t, t+j)$, where the latter stands for future price with delivery at time $t+j$ signed at time t. They can also buy or sell riskless bonds with different maturity. Hence, the speculator's problem conditional on information in moment j is:

$$\underset{X(j),a_{j,t},b(j,t)}{Max} \quad E_j \left[\begin{array}{c} \int_j^T w_{j,t} \left[S(t)X(t)a_{j,t} + F(j,t)X(t)(1-a_{j,t}) + b(j,t) \right] dt \\ - \int_j^T q(j,t)b(j,t)dt - S(j)X(j) \end{array} \right]$$

$$\text{s.t.} \qquad X(t) = \exp(-\delta(t-j))X(j) \qquad 0 \leq a_{j,t} \leq 1 \qquad X(j) \geq 0 \qquad t > j$$

Where $w_{j,t}$ is the stochastic discount factor, $X(j)$ is the amount of goods bought in time j and stored, δ is the depreciation rate (assumed to be zero), $1 - a_{j,t}$ is the fraction of goods that are going to be sold at a futures prices $F(j,t)$ with delivery t signed in moment j, and $b(j,t)$ are risk free bonds sold at moment j at price $q(j,t)$ with maturity $t \geq j$.

In equilibrum, the agent must be indifferent, in order to avoid arbitrage possibilities:

$$S(j) \geq q(j,t)F(j,t) = E_j(q(j,t)S_t) = \hat{E}_j(S_t) \qquad 0 \leq t \leq T$$

under some equivalent measure

17.5 The goal: correlation under backwardation

Since the goal of the paper is to explain correlation under backwardation cases (with contango the answer is trivial), I will assume that expectations regarding the harvest's size are such that contango cases are very unlikely or have close to zero probability, that is, the market assumes that we will be under backwardation almost surely. This will make results more powerfull.

Under backwardation, we have a pair A^*, B^* that depends on the size of H. For ever pair A^*, B^* there exists an optimal consumption path $x^*(t)$.Hence, $x(.)$ depends also on random variable H.Reminding that for a decentralized equilibrium, prices are given by:

$S(t) = D^{-1}(x^*(t))$

We can take expectations for some measure to obtain future price curve $F(t, t + j)$

$$F(0, t + j) = \hat{E}_0(S(t))$$

$$\begin{aligned} F(0,t) &= \underbrace{\exp(-rt)Z\left[\frac{\frac{r}{\rho}}{\left[1-\exp(-\frac{r}{\rho}(A^*-J)\right]}\right]^{-\rho}}_{M(t,,A^*-J)} \hat{E}_0\left(\left[G(A^*)H + \bar{Q}^*\right]^{-\rho}\right) \\ J &\leq t \leq A^* \end{aligned} \tag{10.a.2}$$

$$\begin{aligned} F(0,t) &= Z\hat{E}_0\left[g(t)H\right]^{-\rho} \\ A^* &< t \leq B^* \end{aligned} \tag{10.b}$$

$$\begin{aligned} F(0,t) &= \underbrace{\exp(-r(t-B^*)Z\left[\frac{\frac{r}{\rho}}{\left[1-\exp(-\frac{r}{\rho}(T-B^*)\right]}\right]^{-\rho}}_{M(t-B^*,T-B^*)} \hat{E}_0\left(\left[(1-G(B))H\right]^{-\rho}\right) \\ B^* &\leq t \leq T \end{aligned} \tag{10.c}$$

And for delivery dates before J

$$\begin{aligned} F(0,t) &= \underbrace{\exp(-rt)Z\left[\frac{\frac{r}{\rho}}{\left[1-\exp(-\frac{r}{\rho}J)\right]}\right]^{-\rho}}_{M(t,J)} \left[Q(0)-\bar{Q}^*\right]^{-\rho} \\ 0 &\leq t \leq J \end{aligned} \tag{10.a.1}$$

Hence, the speculator will be indifferent between buying, selling and storing physical goods. For each moment t he desires has some level of storage $X(t)$. Therefore, the law of motion for his demand will be given by the market clearing equation:

$$\dot{X}(t) + c(t) = y(t) \tag{C.2}$$

and

$$\dot{Q} = \dot{X}$$

Since there must be incentives to have $\bar{Q} > 0$,

$F(0, J^-) = F(0, J)$;

Hence

$$\exp(-rJ)\left[\frac{1}{\left[1-\exp(-\frac{r}{\rho}J)\right]}\right]^{-\rho}\left[Q(0) - \bar{Q}^*\right]^{-\rho} = \left[\frac{1}{1-\exp(-\frac{r}{\rho}(A^*-J))}\right]^{-\rho} E_0\left[\left(G(A^*)H + \bar{Q}^*\right)^{-\rho}\right]$$

Which is satisfied by condition B.6.A (solution for backwardation).