Financial and Commodity-specific expectations in soybean futures markets

A.N.Q. Cao; S.-C. Grosche

Institute for Food and Resource Economics, Rheinische Friedrich-Wilhelms- Universität Bonn, , Germany

Corresponding author email: an.cao@ilr.uni-bonn.de

Abstract:

We conceptualize the futures price of an agricultural commodity as an aggregate expectation for the spot price of a commodity. The market agents have divergent opinions about the price development and the price drivers, which initiates trading. In informationally efficient markets, the price will thus reflect expectations about its influencing variables. Using historical decompositions from an SVARX model, we analyze the contribution of financial and commodity- specific expectation shocks to changes in a trading-volume weighted price index for corn and soybean futures on the Chicago Board of Trade (CBOT) over the time period 2005- 2016. Financial expectations are instrumented with the DJ REIT Index, commodity demand expectations with the CNY/USD exchange rate and supply expectations with changes in the vapor pressure deficit. Results show that the price index was affected by cumulative shocks in the REIT index during the time of the food price crisis, but these shocks are only of small magnitude. Weather fluctuations have a minimal impact on the week-to-week fluctuation of the commodity price index.

Acknowledegment:

JEL Codes: C32, C52

#1831

Financial and commodity-specific expectations in soybean futures markets[1]

Abstract. We conceptualize the futures price of an agricultural commodity as an aggregate expectation for the spot price of a commodity. The market agents have divergent opinions about the price development and the price drivers, which initiates trading. In informationally efficient markets, the price will thus reflect expectations about its influencing variables. Using historical decompositions from an SVARX model, we analyze the contribution of financial and commodity- specific expectation shocks to changes in a trading-volume weighted price index for corn and soybean futures on the Chicago Board of Trade (CBOT) over the time period 2005- 2016. Financial expectations are instrumented with the DJ REIT Index, commodity demand expectations with the CNY/USD exchange rate and supply expectations with changes in the vapor pressure deficit. Results show that the price index was affected by cumulative shocks in the REIT index during the time of the food price crisis, but these shocks are only of small magnitude. Weather fluctuations have a minimal impact on the week-to-week fluctuation of the commodity price index.

JEL classification: C32, C52, G15, Q02

Key words: Structural Vector Autoregressive models with exogenous variables (SVARX); commodity markets; financialization; food price crisis

1 Introduction

Commodity market phenomena of the more recent past include peaking and plummeting prices and a surge in trading volume. The “food price crisis” in 2007/08 renewed interest in investigating the (short-run) price drivers of various agricultural commodity markets. But, as of today there is still a lack of consensus on the specific contribution of shocks to market fundamentals, especially weather patterns, and shocks resulting from a financialization of commodities. A related question is whether these patterns were different during and after the food price crisis compared to the years before.

Structural simultaneous equation models are a good candidate to model the joint development and interaction of several endogenous variables. Such models have in the past been frequently applied in macroeconometrics. Recently, many methodological advances have been made in the subgroup of structural vector autoregressive models (SVARs). First proposed by Sims (1980), SVARs have since been further developed with a focus on addressing the challenge of finding credible identifying restrictions on the error covariance matrix. In the standard recursively identified models with short-run restrictions, emphasis is put on variable selection and careful economic argumentation to avoid the critique of atheoretical and therefore

[1] An earlier version of this paper was presented at the 15th EAAE congress in Parma, August 29 - September 1 2017

incredible restrictions that the early SVAR literature had to face (e.g. (Kilian 2009)). SVAR models, under the name of Structural VARX or SVARX (Kilian and Lütkepohl 2016; Lütkepohl 2005; Ocampo and Rodríguez 2012, 2012), also permit the inclusion of fully exogenous variables.

The objective of this paper is to assess the contribution of financial as well as commodity demand- and supply-specific expectation shocks to changes in the futures prices for soybeans on the Chicago Board of Trade (CBOT) over the weeks from 01/2005-11/2016. We choose a recursively identified SVARX model with short-run restrictions on the variance-covariance matrix of error terms. We put emphasis on the selection of variables that conform to the underlying economic model and instrument both commodity-specific and financial expectations with suitable variables. Historical decompositions then allow for assessing the response of one variable to cumulative shocks in another variable up to a specific point in time (Kilian and Lütkepohl 2016, pp. 117).

To help variable selection and ultimately develop credible restrictions, we limit the market focus. The soybean market is particularly interesting due to its global importance and its unique demand side situation with 63% of global soybean imports going to China. On the other hand, from a financial perspective, soybeans have relatively high trading volumes on the Chicago Board of Trade (CBOT) and are included in important commodity indices such as the S&P GSCI. This paper contributes to the existing literature as there is a need to further explore the potential of SVAR framework to investigate the contribution of structural shocks in the development of prices for important agricultural commodities. And, historical decompositions, despite their advantage in assessing cumulative shock contributions to the development of one variable over time, are still underrepresented in present structural analyses. We also add another perspective to the discussion on the potential influence of weather variables and financialization effects on the price dynamics in agricultural futures markets.

The remainder of the paper is structured as follows. The next section describes some theoretical assumptions underlying the model and the background of soybean market, both from fundamental and financial perspectives. In section 3 we present the economic concepts behind the variable choice and give an overview of SVARX methodology. Section 4 presents the empirical application and its results, section 5 concludes and gives an outlook on planned future extensions.

2 Conceptual framework and market background

This section describes the foundation of our theoretical model. Later, we also provide a brief overview of soybean market both from a fundamental and a financial perspective.

2.1 Conceptual framework

Conceptually, the futures market is postulated to be a market for expectations of the future spot price. Agents exchange information on their price expectations by trading contracts. Thus, contract volume is a proxy for information influx into the market. Market agents are assumed to have heterogeneous priors and use different information sets. Heterogeneous priors as proposed by Morris (1995) imply that agents process even the same information differently. Thus, differences in opinion about the expected spot price do not automatically stem from different information sets, but may be a result of differences in agents' prior beliefs concerning how to interpret a given information signal. This results in a "difference-in-opinion" (DO) environment, where traders disagree with each other even when they have the same information. DO models can better explain soaring trading volumes, as experienced on agricultural commodity markets (Figure 1) – without having to resort to stochastic "noise traders" (Banerjee and Kremer 2010; Morris 1995). Models in this spirit have been applied in the financial market literature (e.g. Banerjee and Kremer (2010)) and, recently, also to derive market equilibria under hedging and speculation on agricultural futures markets (Fishe et al. 2014).

A typical example for differences in priors is expert versus non-expert traders. On the agricultural commodity markets, traditional market participants such as producers, processors or big commodity trading houses all have exposure to the physical commodity and are typically experts for specific commodity markets. Financial traders, on the other hand, have no exposure to the physical commodity and rather trade on the joint development of commodities in an index or on expected portfolio correlations. Even when all traders have access to the same (public) information, based on their status, they may interpret it differently when deciding on their trading activity. In informationally efficient markets, all information will be reflected in the price. Thus, c.p. an increased presence of financial traders should make the futures price more susceptible to shocks in variables that only these traders consider important enough to change the price.

To determine whether the futures price is determined rather by commodity- specific fundamental shocks or by financial shocks, variable selection is crucial. An important point to consider is that agents will not only formulate expectations about the development of the futures price, but also about the development of all variables that they include in their information set to determine the futures price, based on their prior (see Grosche (2014) for related arguments). Thus, in our empirical model we attempt to select variables that reflect expectations about a commodity-specific variable or a financial variable.

As we seek to model short-term effects, we also select our model variables based on their importance for short-term (week-to-week) price fluctuations. On the demand side, we seek to instrument short-term shocks with the exchange rate between the Chinese Renminbi and the US Dollar. On the supply side, an important driver behind short-term variations in yield is the weather during the growing season, i.e. precipitation and temperature. While the exact relationship behind weather, technology and crop yields is complex, weather expectations can be assumed to influence agents' expectations of yield and therefore also expectations about the future spot price. An example for an explicit incorporation of weather variables in econometric models for crop yields and prices is Roberts and Schlenker (2013) who use weather variables to instrument exogenous price effects in a structural model for supply and demand of four basic staple crops (corn, rice, soybeans, wheat). For the financial expectations, we seek a variable that has no or little fundamental connection to the agricultural commodity markets, but may reflect relevant financial expectations.

2.2 Market background

Fundamental perspective

Soybeans are C3 oilseed crops, mostly used as animal protein feed and vegetable oil. The soybean market is strongly concentrated, both in terms of main producers/exporters and consumers/importers. On the supply side, according to USDA-FAS (2017), the top three producers (United States (US), Brazil, Argentina) produced a combined quantity of 280,708 thousand metric tons in the year 2016/17, which is equivalent to 82.4% of global production. These countries also dominate exports by providing an equivalent of 88.7% of global exports. The number-one export destination for all countries is China, accounting for 57% of US, 68% of Brazilian, and 81% of Argentinian soybean exports in the period 2008/09 - 2012/13 (data taken from Global Trade Atlas cited in Meade et al. (2016)). In the year 2016/17, Chinese soybean imports account for 63% of total imports, followed by the EU (10%) (USDA-FAS 2017). In the US, soybean production is concentrated in the center, mostly in the Heartland region. It encompasses Indiana, Illinois, Iowa, Minnesota, Missouri, Nebraska, South Dakota, Kentucky and Ohio. The region is characterized by fertile, well-drained soil and moderate climate (Meade et al. 2016). According to USDA data, the Heartland states account for a combined share of 73% of US soybean production, on 2010-14 averages. Illinois and Iowa are the two largest contributors to national soybean production with a share of 14% each (USDA 2017c).

Most of the soybean demand is for crushing in order to produce soybean oil or soybean meal, which is then used e.g. as animal feed or for bioenergy. In 2016/17, out of the total production of about 341 m tons, 291.5 m tons are for crushing (86%). In the year 2016/17, Chinese crushing accounts for about 30% of total soybean crushing, followed by the US (18%), Argentina (15.5%), Brazil (14%) and the EU (5%). These countries together account for 82% of total crushing. Consequently, China is by far the biggest importer of soybeans for crushing, as all other major crushers are also major producers of soybeans.

There is a fundamental connection between soybeans and corn. Both are used for animal feed and biofuels (corn for bioethanol and soybeans for biodiesel) and both require similar climatic conditions, even though corn is a C4 crop originating from a tropical environment (Hollinger and Angel 2009). In the US, both crops compete for land as both are grown in the Heartland region and have similar crop calendars. Soybeans are usually planted between May and mid-June, corn is planted between April and May and both crops are harvested between mid-September and early November (USDA 2017a,b). Producers in the Heartland region often grow both crops and have the opportunity to choose between soybean and corn plantings based on their price expectations and weather conditions by planting season. For example, Miao et al. (2015) note that in some years, corn was replaced by increases in soybean acreage in Midwest US due to excessive precipitation in spring. There is also a discussion that the US biofuel mandates initially triggered substitution away from soybeans in favor of corn. A case in point is in 2007/08, a 22% increase in the corn area in the US was accompanied by a 14% reduction in the soybean area (Headey 2011).

Financial perspective

In the years during and after the “food price crisis”, there has been an ongoing debate concerning the effects of “financialization” of commodities. While there is no formal definition, this term typically refers to the joint occurrence of the phenomena of an increased market presence of financial investors on commodity markets on the search for portfolio-diversifying assets (cp. Fortenbery and Hauser (1990)), the creation of new commodity-linked financial investment products and an increased trading activity on the derivative, especially futures, markets (cp. Silvennoinen and Thorp (2013)). While most of the financial investment in agricultural commodity markets is linked to the investment into broader commodity indices such as the S&P Goldman Sachs Commodity Index (S&P GSCI), replication of these indices will ultimately require some trading activity on the futures markets (Grosche 2014).

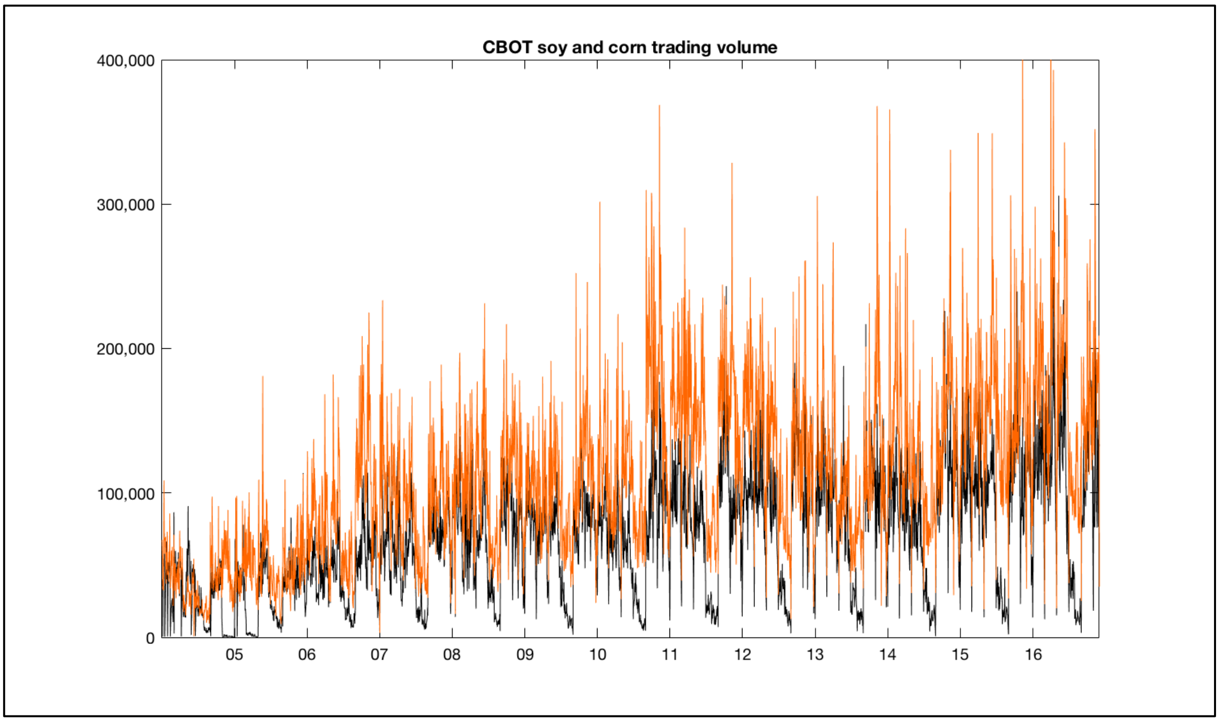

Figure 1: Trading volume for CBOT corn and soybean active contracts

Notes: Orange line represents CBOT corn contract and black line represents CBOT soybean contract trading volume

Source: Bloomberg

On the CBOT, the corn contract is the most liquid agricultural contract in terms of trading volume, soybeans and wheat have lower daily volumes, but are still liquid futures markets. Figure 1 shows the development of daily contracts traded for both the corn and soybean contract over the time period 01/2005-11/2016.

3 The model

In this section we present the motivation behind our variable choices. We also provide an overview of SVARX models and their application in the investigation of price drivers on agricultural markets

3.1 An overview of SVARX model

The development path of SVAR methodology

Dynamic Simultaneous Equation Models (DSEM) used to be predominant in macroeconometrics several decades ago. However, since the critique of Sims (1980) toward "incredible restrictions" which are prevailing in macroeconomics, the use of DSEMs has been declined, which gave rise to the classed of VAR models proposed by Sims. The core idea of the VAR approach is that all the variables in the system can be treated jointly endogenous, and thus there is no need for finding reliable exogenous variables for identification. VAR models have faced various critiques for their atheoretical approach with which structural shocks are identified. In this fashion, Choleski decomposition is applied on the variance-covariance matrix of reduced-form residuals to orthogonalize the structural shocks without referring to the economic relationships among variables (Keating 1990). This has motivated the development of structural VAR models (SVARs), in which the focus is to find identifying restrictions that are connected to economic theories. A direct consequence from this development is that apart from the traditional short-run, recursive identification, there exist now many identification methods. For instance, non-recursive or overidentifying restrictions can be imposed to achieve theoretical consistency. It can also be that the model is not point-identified, but rather partly-identified or set-identified, using inequality restrictions. Alternatively, restrictions might be placed on long-run instead of short-run responses, as proposed by Blanchard and Quah (1989). An excellent, thorough overview of these approaches is found in Kilian and Lütkepohl (2016). Nevertheless, to the scope of this study, the extension of SVAR model to incorporate exogenous variables is the most relevant, resulting in the class of SVARX models. As it will become clear below, this model has exactly the same set up as DSEMs.

As an extension of VAR models, $VARX(p,s)$ models still consist of a vector of K endogenous variables y_t, plus a vector of M exogenous variables x_t. The structural form can be written as:

$$A_0 y_t = c^* + A_1^* y_{t-1} + \ldots + A_p^* y_{t-p} + B_0^* x_t + B_1^* x_{t-1} + \ldots + B_s^* x_{t-s} + w_t\,, \tag{1}$$

where p and s are the numbers of lagged terms included in the model of y_t and x_t, respectively. c^* is K-dimensional vector of intercept terms. A_i^*, $i = 1,\ldots,p$ is a $K \times K$ matrix of autoregressive slope coefficients corresponding to lagged endogenous variables y_{t-i}. A_0 thereby collects the coefficients that depict the instantaneous effects among endogenous variables. Likewise, B_i^*, $i = 0,1,\ldots,s$ is a $K \times M$ matrix of autoregressive slope coefficients corresponding to current and lagged endogenous variables x_{t-i}. The K-dimensional vector of white-noise error terms w_t contains mutually uncorrelated elements such that they can be interpreted as structural shocks (Kilian and Lütkepohl 2016, pp.171). Consequently, a structural VARX model can be viewed as a DSEM (see Lütkepohl (2005, pp. 388)).

The methodology

The reduced form of model (1) in lag-operator notation is given by equation 10.2.7, Lütkepohl (2005, pp. 391):

$$A(L)y_t = c + B(L)x_t + u_t, \tag{2}$$

where $A(L) = I_K - A_1 L - \ldots - A_p L^p$ and $B(L) = B_0 + B_1 L + \ldots + B_s L^s$. $c = A_0^{-1}c^*$ is the reduced-form intercept vector. Similarly, the matrices of reduced-form coefficients A_i and B_i are obtained by multiplying A_0^{-1} with their corresponding structural coefficient matrices A_i^* and B_i^*. u_t is the vector of reduced-form residual with variance-covariance matrix Σ_u. This model can be estimated using Restricted Generalized Least Squared methods (see Lütkepohl (2005, section 10.3)), Maximum Likelihood (Kilian and Lütkepohl 2016, section 2.3.4) or Bayesian estimation (Ocampo and Rodríguez 2012).

If all variables in the system are stationary processes which implies that $A(L)$ is invertible, it is possible to obtain the final-form model:

$$y_t = \mu + D(L)x_t + A(L)^{-1}u_t, \tag{3}$$

where $\mu = A(L)^{-1}c$ is the vector of deterministic terms and $D(L) = A(L)^{-1}B(L) = \sum_{i=0}^{\infty} D_i L^i$ consists of $(K \times M)$ dynamic multiplier matrices D_i. The elements of these matrices are interpreted as the marginal effects of the exogenous variables on endogenous variables, in an analogous way to the impulse response matrices of endogenous variables (i.e. those resulting matrices in $A(L)^{-1}$).

Finally, it should be noted that the relationship between w_t and u_t remains the same as in SVAR models, which is: $u_t = A_0^{-1}w_t$. Hence, assuming that we can achieve identification for a unique matrix A_0, it must be the case that:

$$y_t = \mu + D(L)x_t + A(L)^{-1} A_0^{-1} w_t . \tag{4}$$

As all the elements of this equation are solved, the effects of both structural shocks in those endogenous variables and exogenous variables on the system become clear.

To facilitate the historical decomposition that we use in the next section, we shall proceed from equation (4) the Structural Vector Moving Average (Structural VMAX) representation of the model. Define $\psi(L) = A(L)^{-1} = I_K + \psi_1 L + \psi_2 L^2 + \dots$, it follows directly that:

$$y_t = \mu + \sum_{i=0}^{\infty} D_i x_{t-i} + \sum_{i=0}^{\infty} \psi_i A_0^{-1} w_{t-i} , \tag{5}$$

where ψ_i, $i = 1, 2, \dots, \infty$ is the matrix of VMAX coefficient at lag i, taken from the polynomial $\psi(L)$. Using the same argument in Kilian and Lütkepohl (2016, pp.114), for a stationary VARX, equation (5) can be approximated to:

$$y_t = \mu + \sum_{i=0}^{t-1} D_i x_{t-i} + \sum_{i=0}^{t-1} \psi_i A_0^{-1} w_{t-i} . \tag{6}$$

SVAR application to investigate agricultural commodity price drivers

As said before, much of the work in applied SVAR modelling has been done in the area of macroeconometrics. For commodity price drivers, most works concern crude oil market. Examples include Kilian (2009) and Kilian and Murphy (2014). It was primarily in the aftermath of the food price crisis that SVAR models also started to be more frequently applied to investigate the structural relations between agricultural commodity prices and fundamental and financial variables. Some applications of this model include McPhail et al. (2012), Enders and Holt (2014), Baumeister and Kilian (2014) and Hausman et al. (2012). On the contrary, the use of SVARX is limited among applied VAR literatures. In macroeconomic it is due to the difficulty to find "true" exogenous variables (Kilian and Lütkepohl 2016, pp. 72). In the field of agricultural economics, Gutierrez et al. (2015) is one of the very rare VARX applications, in which country-specific VARX models are combined to a global VAR (GVAR) model to investigate price drivers on the global wheat market.

3.2 Variable selection

Corn-Soybean trading-volume-weighted price index

Due to the connections between the soybean and corn markets from a fundamental and financial perspective, it is useful to consider both futures prices jointly. However, as we seek to identify our structural model recursively, we do not consider both price series, but rather combine them into an index. As index weights, we use the annual average trading volume of the futures contracts. An advantage of this trading-volume-weighted price index is that it simultaneously considers the price development and the year-to-year development of trading volume, i.e. a proxy for the disagreement of traders and potential information influx. The index is computed as:

$$P_{Index,t} = w_{S,T-1} P_{S,t} + w_{C,T-1} P_{C,t}\,, \tag{7}$$

where $T = 1, ..., 12$ for the years from 2004 to 2015 and $t = 1, ..., 622$ for the weekly price observations. $w_{S,T-1}$ and $w_{C,T-1}$ denote the trading volume weights of soybean and corn, respectively, calculated based on the trading volume data of the year before[2]. $P_{S,t}$ and $P_{C,t}$ are weekly observations of corn and soybean closing prices on a certain weekday.

CNY-USD exchange rate as a proxy for commodity-specific demand expectations

As most of the soybeans are crushed before they are marketed, short-term demand side shocks will primarily stem from the expected demand of the crushing mills. As said, as of today China accounts for about 30% of global crushing and is the only major soybean crushing country that mostly crushes imported soybeans. Short-term effects on soybean demand for crushing may thus stem from the onshore exchange rate between the Chinese Renminbi (CNY) and the US Dollar (USD). For crushing mills the crush spread is a trading strategy that entails a long position in soybean futures contracts and short positions in soybean oil and meal contracts. The crush spread is the difference between the value of the soybeans contract and the combined value of the soybean oil (BO) and meal (SM) contracts, which measures the profit margin of the crushing mill. In percentage terms, each unit of soybeans traded on the CBOT produces 80% meal, 18.3% oil and 1.7% waste. For a US crushing mill, the crush spread would be calculated as

$$0.8 \times P_{SM} + 0.183 \times P_{BO} - P_S\,, \tag{8}$$

where P_{SM}, P_{BO}, and P_S are the prices for CBOT soybean meal, soybean oil and soybean futures, respectively.

A Chinese crushing mill would consider the US soybean contract but then use the Dalian Commodity Exchange (DCE) soybean oil and meal contracts and thus, apart from considering

[2] It is important to note that the weights are calculated based on the yearly-averaged trading volume of the before, because the volume on the exact last trading day of the year is often not reliable – author note.

differences in metric units and contract sizes, most importantly has to take into account the CNY/USD exchange rate (CME 2015). If the Renminbi depreciates against the US Dollar, the crush spread for Chinese crushers becomes smaller, potentially decreasing their soybean demand and also their hedging demand for soybean futures contracts.

Historically, the CNY/USD exchange rate was fixed. In 2003, when the Renminbi appeared clearly undervalued, China was increasingly under pressure to end this fixed exchange rate regime. In 2005, China officially ended the pegging of the Renminbi and allowed it to fluctuate against a basket of currencies, but within a limited band. In 2007, the trading band of the Renminbi was expanded, but then limited again and in 2012 during the financial crisis, and expanded for another time in 2014. Nevertheless, as of today the Renminbi is likely not fully de-pegged.

Daytime vapor pressure deficit (VPD) as a proxy for commodity-specific supply expectations

There are a wide range of options for instrumenting the weather variable which based on minimum and/or maximum temperature, precipitation level, etc. In this first attempt to model weather effect on agricultural markets, we choose VPD as the starting point, due to its widely acknowledged relevance to plant growth rates. VPD measures atmospheric water demand, based on air temperature and humidity, and has been shown to be highly correlated with extreme heat conditions, which may cause serious water stress on plants (see Lobell et al. (2013; 2014) for detailed discussion). Without the presence of direct humidity measure, VPD is calculated as the difference between saturated vapor pressure (es) at minimum and maximum daily temperature: $VPD = es\left(T_{max}\right) - es(T_{min})$, where for a given temperature T:

$$es\left(T\right) = 0.6108 * \exp(17.269 \times T) / (237.3 + T). \tag{9}$$

Real estate price index as a proxy for financial expectations

Commodities are, from a financial perspective, attractive portfolio-diversifying assets. They are perceived to offer protection against inflation and have low or negative correlations with other financial assets (e.g. Ankrim and Hensel (1993); Gorton and Rouwenhorst (2006)). And, like real estate, commodities are real assets. As described in Grosche and Heckelei (2016), tactical portfolio allocation may rebalance portfolio asset weights as a consequence of shocks in selected markets or due to cross-market arbitrage. During the subprime crisis, real estate assets depreciated in value and the markets were drained of liquidity. For financial investors looking for alternative real assets to place their liquidity, commodities would have been a good candidate. Results in Grosche and Heckelei (2016) showed that indeed, volatility interdependence between soybean, corn, wheat markets and the Dow Jones Real Estate Investment Trust Index (DJ REIT) increased in the years during and after the financial crisis. From a fundamental perspective we see no short-term connection between markets for real estate assets and for agricultural commodities. We therefore postulate that real estate prices are a way to approximate financially motivated expectations on commodity markets.

4 Empirical application

For the empirical part, we first describe our dataset and the computation of the selected variables before we proceed with the estimation of the reduced form VARX model, the subsequent derivation of the coefficients for the restricted error covariance matrix for the SVARX model and the computation of historical decompositions. In the final subsection we discuss our results.

4.1 Data

Data for daily corn and soybean futures prices and CNY/USD (on-shore) exchange rates are obtained from Bloomberg for the time period 01/2005-11/2016 and aggregated to weekly level, leading to 622 observations. The prices are closing prices for the active contracts, rolled over on the last trading day (relative to expiration). Trading volume for these contracts is also obtained from Bloomberg, for the time period 01/2004-11/2016. Following Grosche and Heckelei (2016), DJ Equity REIT index is selected to instrument financial expectations. The index reflects the value of all publicly traded real estate investment trusts in the Dow Jones U.S. stock universe (i.e. US publicly traded companies within Down Jones stocks indices that primarily own and operate income-producing real estate, and are classified and taxed as equity REITs). Data on end-of-day index quotes is also obtained for the time period 01/2005-11/2016.

Weather records are collected from USDA Weekly Weather and Crop Bulletin for the period 01/2005-11/2016, which is released regularly on Tuesday by 4:00 p.m ET. Due to this reporting timing, we choose Wednesday as the weekly aggregating point for all price and price index series in the model, in order to capture any possible effect of the weather information contained in the report on the price index. The Bulletin comprises data of temperature and precipitation at 225 stations across 51 states of the US[3]. From this, we extract the data for 47 stations in 9 Heartland states where corn and soybean production is

[3] Data was extracted manually from pdf-formated reports, as none of such combined dataset is available online.

concentrated, namely Iowa, Illinois, Indiana, Kentucky, Minnesota, Missouri, Nebraska, Ohio and South Dakota. At each time t, simple average is used to aggregate minimum and maximum temperature data to state-level, before we compute weighted averages for the whole region. The weight allocated to each state is the simple average of corn and soybean production shares of that state, out of the total shares of the whole Heartland region in total US production. The VPD is then calculated for the whole region based on this final aggregated observation.

To capture the week-to-week fluctuations, all the price indices, exchange rates and also VPD enter the model in form of returns, given by:

$$R_t = ln(P_t / P_{t-1}), \tag{10}$$

leading to 621 observations for all returns series. Figure 2 shows the development of the variables over the time period of observation.

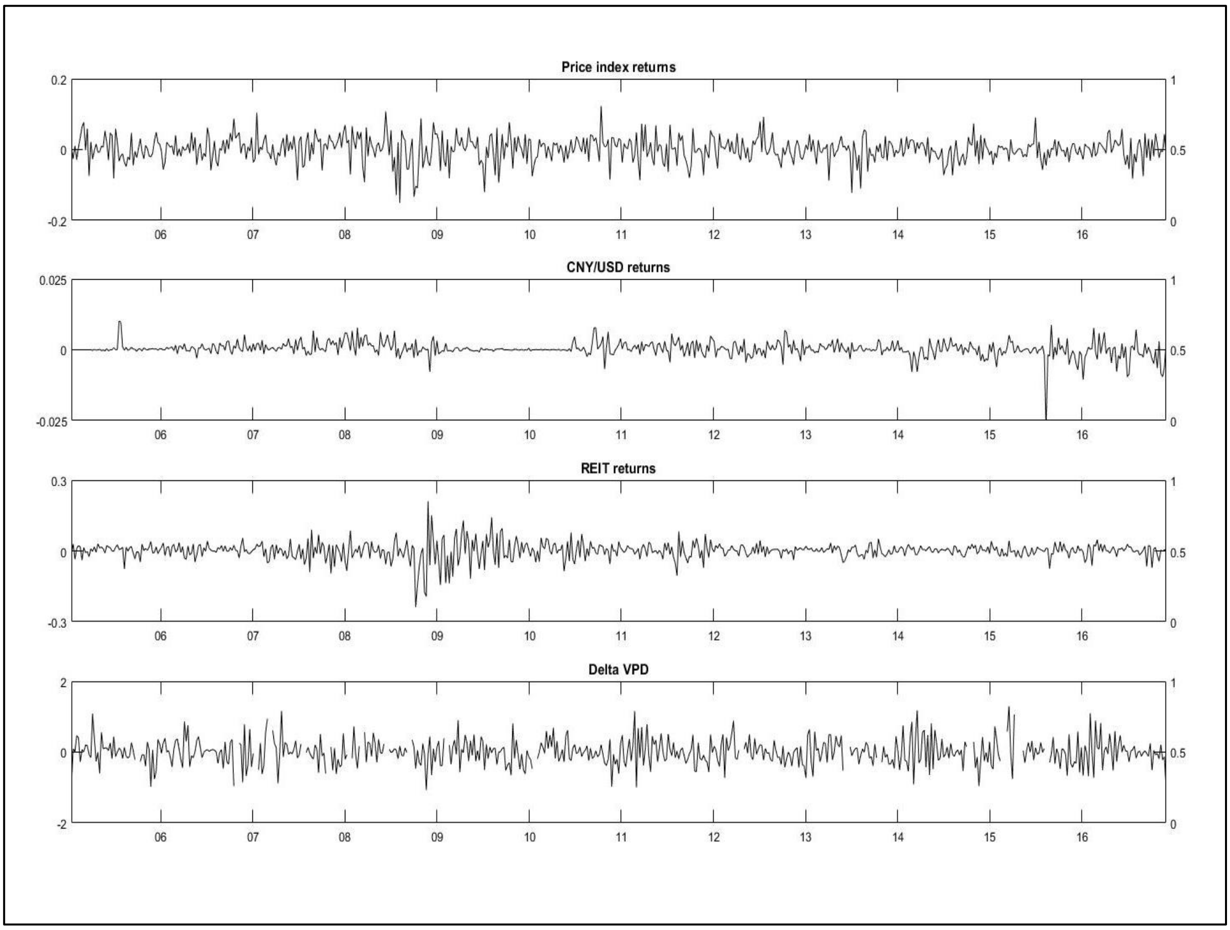

Figure 2: Variables

Source: Own calculation, Matlab 2017a

Augmented Dickey Fuller (ADF) tests

ADF tests for unit roots have been performed to assess the stationarity of the model. For each series, ADF is carried out in two-step procedure: the first step is to select a proper number of lags to include in the testing model using general-to-specific approach, and the second step is to estimate the model with this lag length and evaluate the null hypotheses of unit roots. In the first step, we estimate AR and AR models with constant (drift) term for each lag length $p*$ from 8 to 0, which for weekly series corresponds to about 2 months. For each of these equations, we use T-tests to check whether the last coefficient which is corresponding to lag $p*$ is significant from zero. As soon as the coefficient for the furthest lag in an equation becomes statistically significant from zero with at least 5% of significant level, then this lag length is used for evaluating the null hypotheses in the next step[4]. Using this procedure, the ADF tests reject the null for all variables at a level of significance of at least 5%.[5] Hence we can conclude that our model is stationary, and therefore is valid for stationary SVARX inference presented in previous section.

4.2 Model estimation and identification

In a theory-oriented approach, we apply a fixed lag order selection based on knowledge about the markets as suggested by Kilian and Lütkepohl (2016, pp.60), rather than data-dependent procedures. We choose to include 8 lags – corresponding to 2 months – of endogenous variables y_t in the model, based on the observation that financial and commodity futures markets both can be assumed to be informationally efficient, with respect to short-term movements of relevant variables, and hence longer lag length would be implausible. To maintain generality in this first model, the same lag number is applied for x_t (i.e. $p = s = 8)$, though given the stochastic nature of weather, a shorter lag length can be considered.

Without loss of economic grounds, the choice of variables described in the previous section allows us to impose a Wold causal chain among the endogenous variables, which results in lower-triangular matrix of instantaneous effects A_0. The resulting vector y_t and x_t are:

$$y_t = \begin{bmatrix} \Delta REI \\ \Delta CNY \\ \Delta IND \end{bmatrix}; \qquad \text{and} \qquad x_t = \left[\Delta VPD\right],$$

[4] See Enders (2014, pp. 215 ff.)
[5] Detailed results from the ADF tests are available from the authors on request.

where ΔREI is returns for REIT index, ΔCNY is returns for CNY/USD exchange rates, ΔIND is returns for Corn-Soybean price index and ΔVPD is the VDP difference term.

Specifically, we postulate that REIT index returns is only instantaneously affected by its own shocks (i.e. real estate market shocks), returns on CNY/USD exchange rate can be instantaneously affected by both real estate market shocks and CNY/USD exchange rate shocks but not by shocks to soybean and corn markets. Finally, all those structural shocks in the model can have instantaneous impact on the Corn-Soybean price index. This is plausible because agriculture in general account for a very small share of global economic activities, therefore unlikely to affect real estate and foreign exchange markets contemporaneously.

Furthermore, we argue that on weekly basic, Heartland VPD fluctuations are only relevant for commodity markets, as financial markets are unlikely affected by weather conditions in the short-run. This implies that all the autocorrelation coefficients of x_t and its lags in ΔREI and ΔCNY are restricted to zero.

With these restrictions, restricted Generalized Least Squares is used to estimate a reduced-form VARX model with intercept terms, following the algorithm described in Lütkepohl (2005, section 10.3.1). The inverse of matrix $\hat{A}_0$ is then obtained from the Choleski decomposition of variance-covariance matrix $\hat{\Sigma}_u$ of reduced-form residuals. The estimated expression $\hat{u}_t = \hat{A}_0^{-1}\hat{w}_t$ is given by:

$$\begin{bmatrix} \hat{u}_{\Delta REI,t} \\ \hat{u}_{\Delta CNY,t} \\ \hat{u}_{\Delta IND,t} \end{bmatrix} = \begin{bmatrix} 0.0371 & 0 & 0 \\ 0.0001 & 0.0027 & 0 \\ 0.0075 & 0.0035 & 0.0349 \end{bmatrix} \begin{bmatrix} \hat{w}_{\Delta REI,t} \\ \hat{w}_{\Delta CNY,t} \\ \hat{w}_{\Delta IND,t} \end{bmatrix} \tag{11}$$

4.3 Historical decomposition

Quite often in applied SVAR studies, conclusions are drawn upon structural impulse response functions (IRF) and/or forecast error variance decomposition (FEVD). It should be noted, however, that these measures only report the average effect of each one-time shock separately, which the shock being often normalized to a typical magnitude (e.g. one positive standard deviation). Though, at any given point of time, a certain variable can be affected by various shocks, of different magnitudes and occurring in many points of time in the past (Kilian and Lütkepohl 2016, pp. 117). To investigate the “actualized” role of each shocks in the development of a certain variable during the sample period, it is necessary to cross-multiply these average effects with their corresponding shocks and cumulate them over the given time horizon. Toward this end, we go beyond IRF and FEVD and conduct historical decompositions.

We apply the algorithm described in Lütkepohl (2005, section 10.5.1) to calculate VMAX and structural VMAX coefficient matrices (i.e. $\hat{\psi}_i$ and $\hat{\Phi}_i = \hat{\psi}_i \hat{A}_0^{-1}$) respectively for the whole sample period. Following Kilian and Lütkepohl (2016, section 4.3), the cumulative effect of each structural shocks over the time on y_{3t} (i.e. ΔIND) is then calculated as follows:

$$\hat{y}_{3t}^{(1)} = \sum_{i=0}^{t-1} \hat{\phi}_{31,i} \hat{w}_{1,t-i}$$

$$\hat{y}_{3t}^{(2)} = \sum_{i=0}^{t-1} \hat{\phi}_{32,i} \hat{w}_{2,t-i} \tag{12}$$

$$\hat{y}_{3t}^{(3)} = \sum_{i=0}^{t-1} \hat{\phi}_{33,i} \hat{w}_{3,t-i} \,,$$

where $\hat{\phi}_{jk,i}$ is the estimated response of variable j to shock k at horizon i and $\hat{w}_{k,t}$ is the k$^{\text{th}}$ estimated structural shock at time t.

The cumulative effect of the exogenous variable on the variable of interest is obtained analogously:

$$\hat{y}_{3t}^{(x)} = \sum_{i=0}^{t-1} \hat{d}_{3,i} x_{t-i} \,. \tag{13}$$

Since we only have one exogenous variable, $\hat{d}_{3,i}$ denotes the dynamic multiplier of ΔVPD on ΔIND at horizon i. Note that because VDP variable enters the model in difference form, we can set the reference value for it equal to zero[6]. That said, any change in the from-week-to-week VDP indicator (- i.e. the difference terms deviates from mean zero) may have an impact on the price index returns. Consequently, the historical decomposition of ΔIND is:

$$\hat{y}_{3t} - \mu = \sum_{n=1}^{3} \hat{y}_{3t}^{(n)} + \hat{y}_{3t}^{(x)} \,. \tag{15}$$

The result of this procedure is plotted in figure 3.

[6] For the discussion about reference value of exogenous variable in SVARX historical decomposition, see Ocampo and Rodríguez (2012, section 5.2.2).

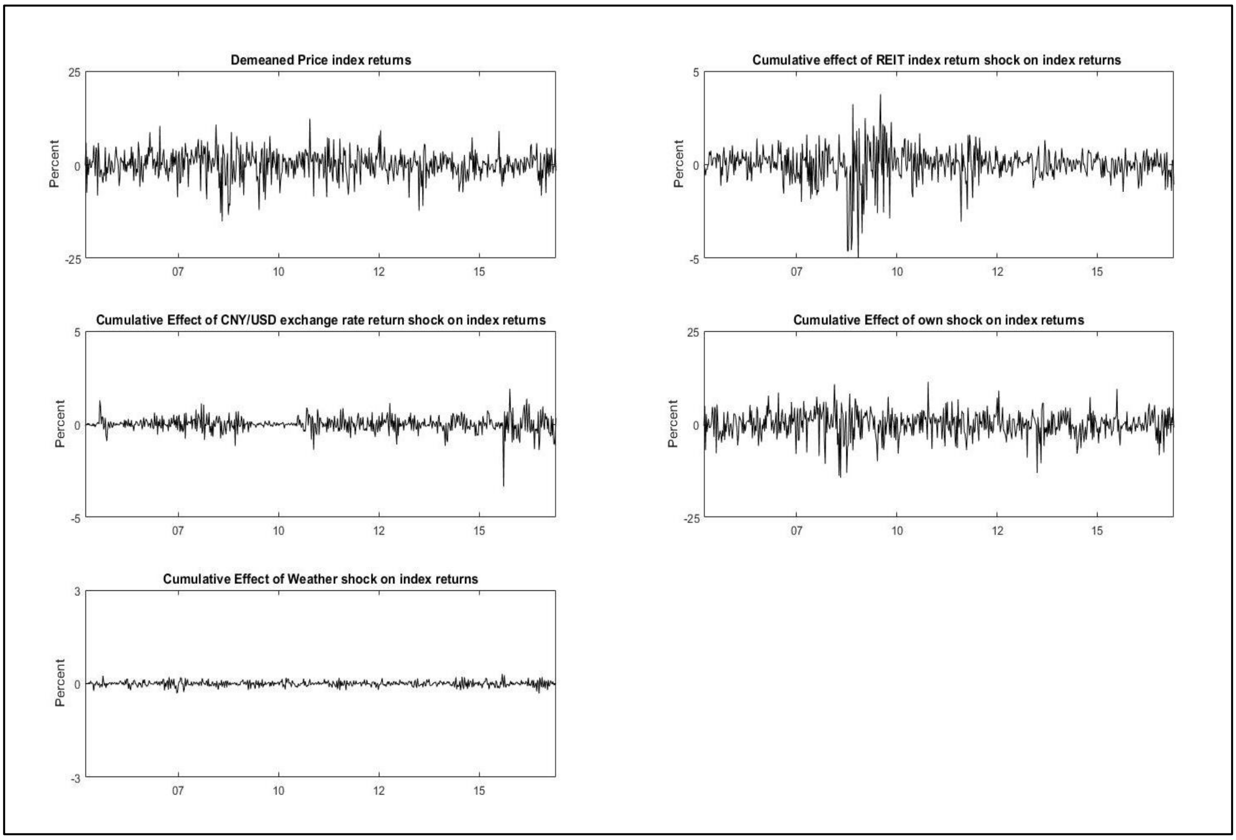

Figure 3: Historical decomposition of Corn-Soybean price index returns during (2005-2016)

Source: Own calculation, Matlab 2017a

4.4 Discussion of results

It is clear that the price index returns are mostly explained by own shocks. In terms of absolute contribution, cumulative shocks to the CNY/USD exchange rate index for the most part explain less than 1% and maximum 3% of the variation in the price index returns. It also appears that cumulative shocks to the REIT index explain at best only about 5% of the price index returns. However, the underlying dynamics do point to an increased response to cumulative REIT index return shocks during the time of the food price and financial crisis, giving some indication that purely financial expectations that had nothing to do with the market fundamentals did play a role in the market dynamics. However, with this short-term focus, even when considering cumulative rather than on-time shocks, there is no indication that these shocks had a marked influence on the absolute level of return variation on the corn and soybean markets. Fluctuations in VDP indicator are shown to have just very minimal effect on the change of the price index. There are two explanations for this. Firstly, it seems that the effect of weather variability is difficult to capture in such short-run models, and thus may need higher level of aggregation. Also, it could be that weather effect on market expectations is non-linear. Secondly, it might be that the effect of weather shocks in form of VDP changes does not enter the information set which are used by traders for their expectation formulation, but rather in another form, such as precipitation or ad-hoc extreme weather events.

Even though with the CNY/USD exchange rate and the DJ REIT index, we have attempted to include variables that will themselves also contain expectations, we cannot exclude that expectations about the future development of these variables are already contained in the futures prices. It is therefore likely that we only capture the response of the price variable to cumulative unexpected shocks in the other variables (see e.g. Kilian (2011) and Grosche (2014) for a discussion of problems related to expectations and forward-looking behavior in the data).

5 Conclusion

In this paper, we have attempted to investigate the role of commodity-specific and purely financial expectations in determining the price dynamics on the futures market for soybeans with help of a recursively estimated SVARX model. Conceptually, we view the futures price as an aggregate of expectations of the future spot price that will incorporate also expectations about potential price drivers. We deviate from the strict rational expectations assumptions and allow market participants to have heterogeneous priors and therefore to disagree and translate their disagreement into trading positions. In the empirical part, the main emphasis was on selecting adequate variables for both the commodity-specific and financial expectations. We have chosen the returns for the CNY/USD exchange rate for the fundamental demand side and VDP returns in the US Heartland region for the fundamental supply side, and returns for the DJ REIT index for the financial side. Due to the linkage of soybean and corn markets, both from a fundamental and a financial perspective, we have decided to analyze these two markets jointly by constructing a trading-volume-weighted price index. Historical decompositions allow for assessing the contribution of cumulative structural shocks and exogenous variables in the development of endogenous variables over time. Compared to IRF and FVED analysis, this tool is relatively underrepresented in the literature on applied structural analysis in SVAR framework. But, they are well suited for our research question.

Results from the historical decomposition show that the price index returns are primarily explained by own shocks. The magnitude of the effects of shocks to the REIT index returns the CNY/USD exchange rate returns or VDP returns is very low. Nevertheless, there is indication that primarily during the food price crisis, financial effects may have played a role in the market dynamics. Due to the informational efficiency of financial markets also with respect to expectations, finding variables that capture new information for the market is, however, challenging.

There are a range of possible strategies for further research. Within this model, given data availability, an idea would be to use the DCE/CBOT soybean crush spread instead of just the CNY-USD exchange rate. That might help addressing the issues of considerable periods with very little variation due to the pegging of the Renminbi to the US Dollar. Another idea would be to construct an exchange rate index taking into account also the Brazilian Real and the Argentinian Peso. Alternatively, one could also try to employ other weather variables that are more intuitively relevant to markets expectations and do not necessarily reflect the real technical relationship with the development of crops. For example, rainfall or extreme degree days can be the candidates. Nevertheless, as said, a week-to-week model is unlikely to capture the weather effect, and therefore should be replaced e.g. by longer time interval, for example monthly level. On the demand side a suggestion is to incorporate expectations on biofuel demand. Moving beyond SVARX model setting, a more complex way to incorporate weather in commodity market models would be to treat weather as a state in which the markets operates rather than a variable and allow for different regimes of model parameters in different states. An example of this is the Markov-Switching approach used in Busse et al. (2012) for analyzing biodiesel supply chain in Germany under changes in market conditions and policy frameworks. In investigate the effect of weather, the seasonal feature of crop production should also be considered, as the effect of weather shocks differs according to their timing with respect to the crop calendar. Another idea is to adjust the price variable to better reflect expectations, for example by adjusting for the risk premium as suggested in Baumeister and Kilian (2014). Very short-term effects of single (extreme) events could be best captured by complementary event-studies. They have the advantage that they can investigate short-term price movements directly around the specific event. Without any power for causal inference, they can nevertheless help to shed light on potential dynamics. Finally, one could follow the idea presented in Banerjee and Kremer (2010) and include trading volume in the model.

References

Ankrim, E. M. and Hensel, C. R. (1993). Commodities in Asset Allocation: A Real-Asset Alternative to Real Estate? *Financial Analysts Journal*, *49*(3), 20–29. https://doi.org/10.2469/faj.v49.n3.20

Banerjee, S. and Kremer, I. (2010). Disagreement and Learning: Dynamic Patterns of Trade. *The Journal of Finance*, *65*(4), 1269–1302. https://doi.org/10.1111/j.1540-6261.2010.01570.x

Baumeister, C. and Kilian, L. (2014). Do Oil Price Increases Cause Higher Food Prices? *Economic Policy*, *29*(80), 691–747. https://doi.org/10.1111/1468-0327.12039

Blanchard, O. J. and Quah, D. (1989). The Dynamic Effects Of Aggregate Demand And Supply Disturba. *The American Economic Review*, *79*(4), 655.

Busse, S. et. alIhle, R. (2012). Price formation in the German biodiesel supply chain: A Markov-switching vector error-correction modeling approach. *Agricultural Economics*, *43*(5), 545–560. https://doi.org/10.1111/j.1574-0862.2012.00602.x

CME. (2015). CBOT Soybeans vs. DCE Soybean Meal and Soybean oil - Crush Spread.

Enders, W. (2014). *Applied Econometric Time Series: 4th Edition* (Fourth edition).

Enders, W. and Holt, M. T. (2014). *The Evolving Relationships between Agricultural and Energy Commodity Prices: A Shifting-Mean Vector Autoregressive Analysis*: University of Chicago Press. Retrieved from http://www.nber.org/chapters/c12810.pdf

Fishe, R. P. H. et. alSmith, A. (2014). Hedging and Speculative Trading in Agricultural Futures Markets. *American Journal of Agricultural Economics*, *96*(2), 542–556. https://doi.org/10.1093/ajae/aat111

Fortenbery, T. R. and Hauser, R. J. (1990). Investment Potential of Agricultural Futures Contracts. *American Journal of Agricultural Economics*, *72*(3), 721. https://doi.org/10.2307/1243042

Gorton, G. and Rouwenhorst, K. G. (2006). Facts and Fantasies about Commodity Futures. *Financial Analysts Journal*, *62*(2), 47–68. https://doi.org/10.2469/faj.v62.n2.4083

Grosche, S.-C. (2014). What Does Granger Causality Prove? A Critical Examination of the Interpretation of Granger Causality Results on Price Effects of Index Trading in Agricultural Commodity Markets. *Journal of Agricultural Economics*, *65*(2), 279–302. https://doi.org/10.1111/1477-9552.12058

Grosche, S.-C. and Heckelei, T. (2016). Directional Volatility Spillovers Between Agricultural, Crude Oil, Real Estate, and Other Financial Markets. In *Food Price Volatility and Its Implications for Food Security and Policy* (pp. 183–205). Springer.

Gutierrez, L. et. alPaolo Roggero, P. (2015). A Global Vector Autoregression Model for the Analysis of Wheat Export Prices. *American Journal of Agricultural Economics*, *97*(5), 1494–1511. https://doi.org/10.1093/ajae/aau103

Hausman, C. et. alBerck, P. (2012). Farm Acreage Shocks and Crop Prices: An SVAR Approach to Understanding the Impacts of Biofuels. *Environmental and Resource Economics*, *53*(1), 117–136. https://doi.org/10.1007/s10640-012-9550-x

Headey, D. (2011). Rethinking the global food crisis: The role of trade shocks. *Food Policy*, *36*(2), 136–146. https://doi.org/10.1016/j.foodpol.2010.10.003

Keating, J. W. (1990). Identifying VAR models under rational expectations. *Journal of Monetary Economics*, *25*(3), 453–476. https://doi.org/10.1016/0304-3932(90)90063-A

Kilian, L. (2009). Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market. *The American Economic Review*, *99*(3), 1053–1069.

Kilian, L. (2011). Structural Vector Autoregressions.

Kilian, L. and Lütkepohl, H. (2016). *Structural Vector Autoregressive Analysis*: Cambridge University Press, forthcoming.

Kilian, L. and Murphy, D. P. (2014). The Role of Inventories and Speculative Trading in The Global Market for Crude Oil. *Journal of Applied Econometrics*, *29*(3), 454–478. https://doi.org/10.1002/jae.2322

Lobell, D. B. et. alSchlenker, W. (2013). The critical role of extreme heat for maize production in the United States. *Nature Climate Change*, *3*(5), 497–501. https://doi.org/10.1038/nclimate1832

Lobell, D. B. et. alHammer, G. L. (2014). Greater sensitivity to drought accompanies maize yield increase in the U.S. Midwest. *Science (New York, N.Y.)*, *344*(6183), 516–519. https://doi.org/10.1126/science.1251423

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*: Springer Science & Business Media.

McPhail, L. L. et. alMuhammad, A. (2012). Disentangling Corn Price Volatility: The Role of Global Demand, Speculation, and Energy. *Journal of Agricultural and Applied Economics*, *44*(03), 401–410. https://doi.org/10.1017/S107407080000050X

Meade, B. et. alDohlman, E. (2016). Corn and Soybean Production Costs and Export Competitiveness in Argentina, Brazil, and the United States.

Miao, R. et. alHuang, H. (2015). Responsiveness of Crop Yield and Acreage to Prices and Climate. *American Journal of Agricultural Economics*, *98*(1), 191–211. https://doi.org/10.1093/ajae/aav025

Morris, S. (1995). The Common Prior Assumption in Economic Theory. *Economics and Philosophy*, *11*(02), 227. https://doi.org/10.1017/S0266267100003382

Ocampo, S. and Rodríguez, N. (2012). An Introductory Review of a Structural VAR-X Estimation and Applications. *Revista Colombiana de Estadística*, *35*(3), 479–508.

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

Silvennoinen, A. and Thorp, S. (2013). Financialization, crisis and commodity correlation dynamics. *Journal of International Financial Markets, Institutions and Money*, *24,* 42–65. https://doi.org/10.1016/j.intfin.2012.11.007

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica*, *48*(1), 1. https://doi.org/10.2307/1912017

USDA. (2017a). Corn map.

USDA. (2017b). Soybean map.

USDA. (2017c). Soybean production overview map.

USDA-FAS. (2017). PS&D (Production, Supply and Distribution) database.