Elsevier Editorial System(tm) for Energy Economics
Manuscript Draft

Manuscript Number: ENEECO-D-13-00402R2

Title: Causality and Predictability in Distribution: the Ethanol-Food Price Relation Revisited

Article Type: Full Length Article

Section/Category: Energy Economics

Keywords: Biofuels; Ethanol; Field Crops; Density Forecasting; Granger Causality; Quantiles

Corresponding Author: Dr. Andrea Bastianin,

Corresponding Author's Institution: University of Milan-Bicocca and Fondazione Eni Enrico Mattei

First Author: Andrea Bastianin

Order of Authors: Andrea Bastianin; Marzio Galeotti; Matteo Manera

Abstract: This paper examines the relationship between biofuels, field crops and cattle prices in the U.S. from a new perspective. We focus on predictability in distribution by asking whether ethanol returns can be used to forecast different parts of field crops and cattle returns distribution, or vice versa. Density forecasts are constructed using Conditional Autoregressive Expectile models estimated with Asymmetric Least Squares. Forecast evaluation relies on quantile-weighed scoring rules, which identify regions of the distribution of interest to the analyst. Results show that both the centre and the left tail of the ethanol returns distribution can be predicted by using field crops returns. On the contrary, there is no evidence that ethanol can be used to forecast any region of the field crops or cattle returns distributions.

ENEECO-D-13-00402R1
Answers to the reviewers' comments

To Prof. Richard S.J. Tol, Editor
Energy Economics

Dear Prof. Tol,

This letter accompanies the revised version of the paper ENEECO-D-13-00402R1: "Causality and Predictability in Distribution: the Ethanol-Food Price Relation Revisited".

We thank you for the consideration of our work and we thank the reviewers for this second round of comments, which have greatly improved the readability of our paper.

As you can see from our detailed reply to the referee's comments, we have addressed all the points raised by reviewer #2.

In addition, we have followed all your suggestions, that is: i) we have shifted as much material as we could to the appendix. Specifically, we have reduced the introduction and the data description, while we have moved part of the comments on the empirical findings to the appendix. With those adjustments, we have been able to reduce the paper length from 31 to 25 pages; ii) we have re-read the paper carefully and checked all the references; iii) we have submitted both the black-and-white and the colour versions of the graphs.

We look forward to your final decision on our paper.

Kind regards,

Andrea Bastianin, Marzio Galeotti and Matteo Manera

ENEECO-D-13-00402R1
Answers to the reviewers' comments

Reviewer #2:

R2/1 *The manuscript is still too long for the amount of information it conveys. Any effort on the part of the authors to reduce the length will be much appreciated by the readers.*

We further reduced the length of the paper by placing the section "Quantile forecasts" in the appendix and by shortening the introduction and the data section. Now the length of the paper, including figures and tables, is 25 pages (the previous version was 31 pages).

R2/2 *As an example, the innovative aspects of the manuscript could be reduced to one paragraph instead of two pages.*

We did as the referee suggests see reply to R2/1.

R2/3 *Please state if prices in figure 1 and table 1 are real or nominal prices. Some statement on how real prices are converted to nominal prices should be provided.*

As described in Section 3 ("Data") we used nominal prices, therefore no conversions were implemented. The title and the notes in Figure 1 have been changed. See R2/5 for Table 1.

Figure 1. Prices: Ethanol, Indices, Field Crops and Cattle

…

Notes: Prices are represented on a common scale (i.e. nominal prices have been multiplied by 100 and divided by their value in January 1987).

R2/4 *Please mention the results of a unit root test on the data. The log differences in prices are stationary, right?*

Results of the Augmented-Dickey-Fuller and Phillips-Perron tests for a unit root, jointly with those of KPSS test for trend-stationarity have now been added to the Appendix. All log prices are integrated of order one, while their first differences are stationary.

R2/5 *Table 1. Tables should stand alone. Please state what panel a and b are within the table.*

Table 1 has been changed as the referee suggests. See below.

Table 1. Descriptive Statistics: January 1987 - December 2010/March 2012

Panel (a): Nominal Prices

	ETH	PI1 (CAT excl.)	PI2 (CAT incl.)	COR	SOY	WHE	CAT
Mean	1.53	102.63	347.25	2.70	6.71	3.84	76.52
Coef. Var.	0.35	0.51	0.28	0.40	0.33	0.38	0.15
Min	0.89	47.46	215.08	1.43	4.00	1.99	58.60
Date Min	01/1987	01/1987	01/1987	02/1987	10/2001	11/1999	09/1998
Max	3.58	287.70	723.11	6.93	13.30	9.84	104.00
Date Max	06/2006	06/2011	12/2010	08/2011	08/2008	03/2008	12/2010

Panel (b): Returns

	ETH	PI1 (CAT excl.)	PI2 (CAT incl.)	COR	SOY	WHE	CAT
Mean	0.09	0.15	0.06	0.09	0.05	0.08	0.19
Coef. Var.	82.05	38.80	55.55	60.25	93.56	66.21	17.67
Skewness	0.40	0.53	0.37	-0.62	-0.23	-0.54	0.02
Kurtosis	4.26	10.76	5.79	6.60	4.81	6.57	4.51

Notes: CAT = cattle; COR = corn; ETH = ethanol; PI1 = price index 1; PI2 = price index 2; SOY = soybean; WHE = wheat. Entries are descriptive statistics for nominal prices (Panel a) and percentage log returns (Panel b). Panel (a) shows the sample average (Mean), the coefficient of variation (Coef. Var.), the minimum (Min) and maximum (Max) price and their dates (Date Min and Date Max). Panel (b) shows the sample average (Mean), the coefficient of variation (Coef. Var.), the sample Skewness and Kurtosis for log-returns. The time period spanned by the monthly nominal spot price of CAT and PI2 is January 1987-December 2010, while the monthly nominal spot prices of COR, ETH, SOY, WHE and PI1 are observed from January 1987 to March 2012.

Research Highlights

- We study the relationship between biofuels, field crops and cattle in the U.S..
- We focus on Granger causality predictability in distribution.
- Density forecasts are constructed using Conditional Autoregressive Expectile models.
- Both the centre and the left tail of the ethanol returns distribution can be predicted by using field crops returns.
- Ethanol cannot be used to forecast any region of the field crops or cattle returns distributions.

Abstract

This paper examines the relationship between biofuels, field crops and cattle prices in the U.S. from a new perspective. We focus on predictability in distribution by asking whether ethanol returns can be used to forecast different parts of field crops and cattle returns distribution, or vice versa. Density forecasts are constructed using Conditional Autoregressive Expectile models estimated with Asymmetric Least Squares. Forecast evaluation relies on quantile-weighed scoring rules, which identify regions of the distribution of interest to the analyst. Results show that both the centre and the left tail of the ethanol returns distribution can be predicted by using field crops returns. On the contrary, there is no evidence that ethanol can be used to forecast any region of the field crops or cattle returns distributions.

Causality and Predictability in Distribution: the Ethanol-Food Price Relation Revisited

Andrea Bastianin[*,†], Marzio Galeotti[‡], Matteo Manera[*]

This Draft: December, 18th 2013
First Draft: July, 5th 2013

Abstract

This paper examines the relationship between biofuels, field crops and cattle prices in the U.S. from a new perspective. We focus on predictability in distribution by asking whether ethanol returns can be used to forecast different parts of field crops and cattle returns distribution, or vice versa. Density forecasts are constructed using Conditional Autoregressive Expectile models estimated with Asymmetric Least Squares. Forecast evaluation relies on quantile-weighed scoring rules, which identify regions of the distribution of interest to the analyst. Results show that both the centre and the left tail of the ethanol returns distribution can be predicted by using field crops returns. On the contrary, there is no evidence that ethanol can be used to forecast any region of the field crops or cattle returns distributions.

Keywords: Biofuels, Ethanol, Field Crops, Density Forecasting, Granger Causality, Quantiles.
JEL Codes: C22, C53, Q13, Q42, Q47.

[*] University of Milan-Bicocca and Fondazione Eni Enrico Mattei, Milan, Italy.
[†] Corresponding Author: Andrea Bastianin, DEMS - Department of Economics, Management and Statistics, University of Milan-Bicocca, Via Bicocca degli Arcimboldi, 8, I-20126, Milan, Italy. E-mail: andrea.bastianin@unimib.it. Phone: +39 02 64455879.
[‡] University of Milan and IEFE-Bocconi, Milan, Italy.

Acknowledgements: The authors wish to thank the Editor and two anonymous referees for their valuable comments. The first author gratefully acknowledges the following research grant: "Dote Ricercatori: FSE, Regione Lombardia". Previous versions of this paper have been presented at: the 20th Annual Conference of the European Association of Environmental and Resource Economists, 26-29 June 2013, Toulouse; the 21th Annual Symposium of the Society for Nonlinear Dynamics and Econometrics, 28-29 March 2013, University of Milano-Bicocca; the IAERE First Annual Conference, 8-9 February 2013, University of Ferrara; the Fondazione Eni Enrico Mattei (FEEM), Milan; the Centre for Research on Energy and Environmental Economics and Policy (IEFE), Bocconi University, Milan; DEMM - Department of Economics, Quantitative Methods and Management, University of Milan; DEMS - Department of Economics, Management and Statistics, University of Milan-Bicocca.

Causality and Predictability in Distribution: the Ethanol-Food Price Relation Revisited

1. Introduction

Large world food price increases and huge price volatilities are generally interpreted as problematic for many developing nations, which are compelled to face higher costs to feed large parts of their populations and have to manage the subsequent political instabilities. The level and volatility recently hit by the price of corn are often viewed as the effects of the massive development of biofuels, ethanol in particular [Mercer-Blackman et al. 2008, Mitchell 2008, Parker, 2013, UNCTAD, 2008]. According to the so-called “Food versus Fuel” claim food price inflation is primarily due to the ethanol production boom. This proposition relies on the implicit assumption that, if the amount of arable land is fixed over the short-run, Granger causality runs from ethanol price to corn prices and from corn prices to the price of other corn-based products.

We examine short-run Granger causality relations for the whole distribution of returns on the price of ethanol, field crops and cattle in Nebraska from January 1987 through March 2012.

We focus on in-sample and out-of-sample short-run relations to answer the following questions: *a*) Can lagged returns on ethanol be used to forecast field crops or cattle returns? *b*) Can lagged returns on field crops predict returns on ethanol? *c*) Is the whole distribution of returns predictable? *d*) Or, is predictability limited to some parts of the distribution?

We provide a number of interesting results. In particular, ethanol has no predictive power for field crops and cattle. This finding holds: *i*) in-sample; *ii*) out-of-sample; *iii*) for the whole returns distribution. Moreover, ethanol can be forecasted using lagged returns on field crops. This result has been obtained: *iv*) in-sample; *v*) out-of-sample; *vi*) for the centre and the left tail of the distribution. Finally: *vii*) there is no evidence of predictability in the right tail of the distribution. While results *i*) and *iv*) are in line with most of the related literature (see Section 2), findings *ii*), *iii*), *v*), *vi*) and *vii*) represent new empirical evidence on the biofuels-food price relation.

Many studies have analysed the impact of biofuels on food prices, along two main lines of research (for a comprehensive survey see Zilberman et al., 2013). The first relies on time-series econometrics to analyse the linkages between biofuel and food prices (Serra and Zilberman, 2013 for a survey). The second, by means of simulation- and theory-based methods, deals with the impact of the introduction of biofuels on food prices (Kretschmer and

Peterson, 2010). Time-series studies show that the price of biofuels is positively correlated with the prices of food and fuels, but that the reverse correlation is very weak.

Our paper can be placed in the first strand of the literature. We analyse short-run Granger causality linkages between returns on ethanol, field crops and cattle in the U.S. by considering their whole distribution, rather than focusing on few specific moments such as the mean or the variance. Compared to previous studies about the ethanol-food relation, our approach is innovative in many respects.

First, we test for Granger causality both in-sample and out-of-sample. On the contrary, a common feature to most of the previous empirical literature is to analyse the relationship between biofuel prices and agricultural prices, using only in-sample Granger causality tests, while nothing is said about the out-of-sample performance of the estimated models.

Second, many studies which are surveyed in Section 2 show evidence of in-sample Granger causality running from field crops prices to ethanol prices, but not vice versa. These findings are entirely based on empirical models for the first or second moments of the variables of interest, which ignore the issue of predictability in other parts of the distribution[1]. Since returns are generally non-normal, their distribution can be hardly summarized by the mean. As a consequence, even if there is no evidence of Granger causality in mean, we might still find evidence of predictability in higher moments (see Cenesizoglu and Timmermann, 2008, Granger and Pesaran, 2000, Pesaran and Skouras, 2002 for theoretical motivations of the use of density forecasts).

We extend the previous analyses by using the Asymmetric Least Squares (ALS) estimator of Newey and Powell (1986) to produce forecasts for the whole distributions of returns, as well as for specific areas of the distributions including the first moment.

Third, our density forecasts evaluation, which relies on the quantile scoring rule of Gneiting and Ranjan (2011), is of interest for a variety of forecasts users. A scoring rule is a loss function for density forecasts, which associates a lower score to a better forecast. The quantile scoring rule assigns more weight to the part of distribution (either centre, tails, right or left tail) which is of interest for a forecasts user. For instance, the tails of the distribution are usually the main focus of risk managers (e.g. Value-at-Risk), while policy makers, who often exploit confidence intervals around point forecasts to assess the effects of potential economic interventions, are generally more interested in the centre of the distribution. For

[1] A notable exception is Nazlioglu (2011), who studies non-linear Granger causality linkages between oil and agricultural commodity prices.

example, the Energy Information Administration has been publishing since 2009 confidence intervals for crude oil and natural gas futures prices in its Short Term Energy Outlook.

Fourth, we include cattle price in the analysis to assess whether biofuel induced price variations, if any, are transmitted along the food market chain. Interestingly, although Serra and Zilberman (2013) identify this as a topic that should be high on the biofuels research agenda, in their survey of the econometric literature there are only three studies that include meat (Balcombe, 2011; Esmaeili and Shokoohi, 2010; Nazlioglu and Soytas, 2011).

The rest of the paper is organized as follows. In Section 2 we briefly review the relevant literature. Section 3 illustrates the data, while in the Section 4 we detail our modelling approach. Section 5 contains the empirical results, and Section 6 concludes.

2. Related Literature

The relation between ethanol and field crops prices has been discussed in the empirical literature from two main perspectives: the assessment of the presence of long-run relationships between fuel and agricultural prices, and the investigation of existence, as well as the direction, of their Granger causality links. Given the approach followed in our paper, in this section we concentrate on contributions pertaining to the second strand of research, while we address the interested reader to Serra and Zilberman (2013) and Nazlioglu et al. (2013) for exhaustive surveys of the time series literature.

The studies testing the presence and the direction of the relationship between fuel and agricultural prices deal with a variety of empirical methods applied to weekly or monthly spot and futures prices: structural vs. reduced form models, linear vs. non-linear models, statistical vs. econometric methods. In general, this literature has tackled the issue of Granger causality only with in-sample analyses. The majority of the contributions find evidence of Granger causality running from the prices of field crops, corn in particular, to the price of ethanol. This result is robust to the method of analysis, to the sampling frequency and the type of price. Ubilava and Holt (2010) is the only study that focuses on out-of-sample predictability. Using weekly averages of U.S. futures prices for the period October 2006 - June 2009 and a non-linear time series model for corn, the authors conclude that the inclusion of energy prices (oil and ethanol) in the model does not improve corn price forecasts[2].

Kristoufek et al. (2012b, 2013a) rely on weekly price data for the period between November 2003 and February 2011 to analyse relations between biofuels (U.S. ethanol and German

[2] These results are consistent with those in Bastianin, Galeotti and Manera (2013).

biodiesel), their production factors and fossil fuels. The authors show that in the U.S. short-term and medium-term Granger causality linkages run from corn to ethanol, but not vice versa.

These findings are consistent with those of Vacha et al. (2013) who use wavelet coherence analysis to study time and frequency dependent correlations between biofuels, agricultural commodities and fossil fuels. The results show that the price of production factors (U.S. corn and German diesel) lead the price of biofuels (ethanol and biodiesel), but not vice versa.

Zhang et al. (2009) estimate a vector error correction model (VECM) on U.S. weekly data for corn, oil, gasoline, ethanol, and soybean prices over the period March 1989 through December 2007. In the pre-ethanol boom period, 1989-1999, the authors find evidence of Granger causality running from the price of corn to ethanol price, whereas a causality reversal occurs in the boom period, 2000-2007.

Zhang et al. (2007) test whether the limit-price hypothesis can explain pricing patterns in the U.S. ethanol-fuel market by means of a structural vector autoregression (SVAR) model estimated on monthly data from April 1998 to July 2005. The variables included in their SVAR model are corn, ethanol, MTBE (i.e. methyl-tertiarybutyl ether), gasoline prices and MTBE and ethanol quantities. The results indicate that corn prices Granger cause the price of ethanol, but not vice versa.

Zhang et al. (2010) use monthly price data for corn, rice, soybeans, sugar, wheat, ethanol, gasoline, and oil from March 1989 through July 2008 to analyse short and long-run impacts of fuels on agricultural commodities in the U.S.. The authors fail to find any evidence of long-run and short-run Granger causality between fuel and agricultural commodity prices.

Saghaian (2010) analyses pair-wise Granger-causality relations by relying on monthly data on oil, ethanol, corn, soybean, and wheat prices for the period January 1996 - December 2008. The results point to the existence of unidirectional relationships running from soybeans and wheat price series to ethanol, and hence indicate that ethanol does not Granger cause soybeans or wheat price series. Moreover, there seems to be a feedback relationship between corn and ethanol prices. However, the author shows that the evidence of causality is stronger from corn price to the price of ethanol than vice versa; in fact, causality running in the opposite direction is statistically significant only at the 10% significance level.

Serra et al. (2011) fit an exponential smooth transition VECM to monthly U.S. data on ethanol, corn, oil, and gasoline prices from 1990 to 2008. An increase in ethanol prices is found to cause an increase in corn prices. However, they also show that corn price hikes, lead to increases in the price of ethanol. Given that corn production is relatively inelastic, at least in the short-run, an increase in the size of the ethanol market will yield corn price increases that in turn will yield higher ethanol prices.

Kristoufek et al. (2012a) analyse the relationships between the monthly prices of biodiesel, ethanol and related fuels and agricultural commodities (corn, wheat, sugar cane, soybeans, sugar beets). Their results indicate that in the short and medium-term the price of corn Granger-causes the price of ethanol, but that there is no causality running in the opposite direction. Moreover, the authors show that an increase in the price of corn positively affects the price of ethanol and that this effect is relatively short-lived.

These results have been confirmed and extended in a subsequent paper by Kristoufek et al. (2013b). By applying minimal spanning and hierarchical trees to the same set of commodities the authors study the linkages between the prices of biofuels and their producing factors, whether these relations are non-linear and change at different frequencies (weekly and monthly). Their findings show that the price transmission mechanism is non-linear and that both ethanol and biodiesel respond to the price of their producing factors, but not vice versa.

Wixson and Katchova (2012) test causality and asymmetric price transmission in the U.S. with monthly price from January 1995 to December 2010 for the following commodities: soybeans, corn, wheat, oil, and ethanol. They find evidence of unidirectional Granger causality running from returns on corn and soybeans to returns on ethanol.

A different viewpoint is offered by Gilbert (2010), who shows that the 2007–2008 food price increases can be hardly attributed to the growing demand for grains as biofuels feedstocks. Rather than being market-specific, the 2007-2008 price hikes can be more convincingly explained by common factors, such as macroeconomic and monetary shocks propagating to food prices through index-based investment in agricultural derivatives markets.

3. Data

The U.S. is the world's largest producer of corn, at 13 billion bushel per year. Since 2005, an average one-third of corn crop production has been diverted from food and dedicated to ethanol production. The expansion of U.S. biofuels has been induced by a number of distinct energy and environmental policies (see Janda et al. 2012 for an overview). Our dataset comprises five monthly time series of nominal spot prices, namely ethanol, corn, soybeans, wheat and cattle[3], recorded in Nebraska from January 1987 through March 2012 (December 2010 for cattle).

Moreover, to summarize field crops and cattle price dynamics, we also consider two price indices. The first index includes the three field crops prices, while the second index adds the price of cattle to the first price index. Both indices are constructed by assigning time-varying weights based on the dollar value of production to the price of different commodities. [4]

The price of ethanol is measured in dollars per gallon, the prices of field crops (i.e. corn, soybeans and wheat) are denominated in dollars per bushel, while the price of cattle is expressed in dollars per hundredweight. Data sources are the Nebraska Energy Office for the price of ethanol, and the National Agricultural Statistics Service maintained by the U.S. Department of Agriculture (USDA) for the prices of crops and cattle. The dollar value of production of field crops and cattle used to construct the price indices have been sourced from the USDA database.

The price series are shown in Figure 1, while summary statistics are reported in Table 1.

[Figure 1 about here]

[Table 1 about here]

As shown in Figure 1, the price of ethanol has experienced two main phases. The first period, from 1987 to early 2000's, is characterized by price stability and low volatility. In the second period, from the second half of 2000's onwards, volatility is higher and prices have a rollercoaster behaviour. A joint inspection of Figure 1 and Table 1 (Panel a) reveals that the second period started with a price increase culminating at a record price of 3.58 dollars per gallon in June 2006. The price of ethanol had another peak, at 2.9 dollars per gallon, in July 2008, just one month after the implosion of the oil price bubble originated in March 2008 (Phillips and Yu, 2011). Descriptive statistics for percentage price changes (i.e. returns) are shown in Panel (b) of Table 1. As expected, the unconditional distributions of all series is slightly asymmetric and displays different degrees of excess kurtosis.

[3] The economic intuition of why cattle and ethanol might be related can be illustrated with a simple example. Suppose that there is Granger causality running from ethanol returns to corn returns. Since ethanol feedstocks are used in the production of other food products such as meat, there might also be Granger causality running from ethanol to cattle.

[4] More details about the dataset and the construction of indices are provided in the Appendix.

4. Density Forecasts with Expectiles

Our forecasting strategy can be illustrated as follows[5]. The variables of interest are the percentage price changes of ethanol (ETH), corn (COR), soybean (SOY), wheat (WHE), cattle (CAT), and two price indexes (PI1, PI2). Variables are indexed by the subscript i = ETH, COR, SOY, WHE, CAT, PI1, PI2. Percentage price changes (i.e. returns) on each variable are computed as $r_{it} = 100 \times \ln(P_{it} / P_{it\text{-}1})$, where P_{it} is the price of variable i at time t. We are interested in bivariate relations between returns on ethanol and returns on the other variables, therefore we consider the following single-equation expectile models:

$$\tau_{jt}(\omega|\Omega_{\tau-1}) = \beta_0(\omega) + \beta_1(\omega)r_{\mathrm{ETH}t\text{-}1} + \beta_2(\omega)r_{jt\text{-}1} + \beta_3(\omega)|r_{jt\text{-}1}| + \varepsilon_{jt} \qquad (1a)$$

$$\tau_{\mathrm{ETH}t}(\omega|\Omega_{\tau-1}) = \gamma_0(\omega) + \gamma_1(\omega)r_{jt\text{-}1} + \gamma_2(\omega)r_{\mathrm{ETH}t\text{-}1} + \gamma_3(\omega)|r_{\mathrm{ETH}t\text{-}1}| + \varepsilon_{\mathrm{ETH}t} \qquad (1b)$$

where j = COR, SOY, WHE, CAT, PI1, PI2 and $t = 2,\ldots,T$.

In models (1a)-(1b), $\tau_{jt}(\omega\,|\,\Omega_{t\text{-}1})$ and $\tau_{\mathrm{ETH}t}(\omega|\Omega_{\tau-1})$, for $\omega \in (0, 1)$, denote the 100ωth conditional expectile of returns on variable j and ethanol, respectively, while $\Omega_{t\text{-}1}$ is the information set available at time t-1. The absolute value of returns of the dependent variable is introduced to capture time variation in the conditional distribution of returns (Engle and Manganelli, 2004). Expectile models are indexed by ω. As explained below, we match expectiles with quantiles, indexed by $\alpha(\omega)$ = 0.05, 0.10, ,…, 0.95, and estimate a total of 19 equations for each dependent variable.

Models (1a-b) are similar to the Conditional AutoRegressive Expectile (CARE) of Kuan et al. (2009), the only difference being the inclusion of the additional explanatory variables $r_{jt\text{-}1}$ and $|r_{jt\text{-}1}|$ ($r_{\mathrm{ETH}t\text{-}1}$ and $|r_{\mathrm{ETH}t\text{-}1}|$). For this reason, we refer to models (1a)-(1b) as CARE-X.

Each model (1a) and (1b) (i.e. for a total of 12 models) is estimated with Asymmetric Least Squares (ALS; see Newey and Powell, 1987). ALS is similar to Ordinary Least Squares (OLS), with the exception that the squared error loss function is weighted according to the sign of the residuals.

The solution for the ALS estimator is known as “expectile”. Expectiles, as quantiles, can be used to describe the distribution of a random variable. Since expectiles are less immediate to

[5] More details about the forecasting strategy are provided in the working paper version of the paper, which is available on the webpage of the first author.

interpret than quantiles, in this paper we follow Efron (1991) and Granger and Sin (2000) and obtain the quantiles by calculating the proportion of in-sample observations lying below the 100ωth fitted expectile curve. We then use the estimated quantiles in density forecasting to analyse the predictability of the distributions of returns on ethanol and field crops prices.

The performance of CARE-X forecasts is evaluated against the following benchmark models[6]:

$$\tau_{jt}(\omega|\Omega_{\tau-1}) = \beta_0(\omega) + \varepsilon_{jt} \tag{2a}$$

$$\tau_{\mathrm{ETH}t}(\omega|\Omega_{\tau-1}) = \gamma_0(\omega) + \varepsilon_{\mathrm{ETH}t} \tag{2b}$$

We refer to specifications (2a)-(2b) as the Constant Expectile (CE) models. Models (2a)-(2b) imply that, for each variable and expectile, the optimal forecast in $t + 1$ is the estimate of the 100ωth expectile at time t.

In-sample and out-of-sample Granger causality tests can be easily calculated in the context of the CARE-X models (1a)-(1b). Moreover, an out-of-sample Granger causality test requires to compare the forecasting performances of models (1a)-(1b) with the CE models (2a)-(2b). Since models (2a)-(2b) assume that returns are unpredictable, out-of-sample tests of predictability can be carried out by asking which models produce the lowest forecast error loss function.

For each model and quantile we use a rolling window of size T_0 to obtain a vector of H forecasts. We evaluate each forecast by means of the quantile scoring rule (QS) of Gneiting et al. (2011). The QS provides a summary of the model's overall ability to forecast the whole distribution (i.e. across all quantiles). Scoring rules have the same interpretation as loss functions: more accurate density forecasts are associated to lower scores, less accurate density forecasts are associated to higher scores.

The QS assigns the same weight to all forecast errors, independently of their location in the support of the distribution. Should the focus be on the tails (e.g. for risk management purposes) or on the centre of the distribution, it would be more appropriate to associate a higher score to the area of interest. Gneiting et al. (2011) have thus proposed a weighted version of the QS (WQS).

[6] Studying bi-variate relations among ethanol, field crops, cattle and price indices amounts to estimating 12 CARE-X models for each of the 19 quantiles of interest (i.e. $\alpha = 0.05, 0.10, \ldots, 0.90, 0.95$). For each model involving field crops and price index 1 we produce 180 forecasts. For models involving cattle or price index 2 we issue 165 forecasts. Therefore the total number of one-step ahead forecasts is $(180 \times 19 \times 8) + (165 \times 19 \times 4) = 39900$.

The WQS uses a weight function that assigns a higher score to the desired part of the distribution. In addition to the uniform QS that does not use any weight function, we consider four WQSs to focus on the following parts of the distribution: 1) centre; 2) tails; 3) left tail; 4) right tail. Therefore we have used a total of five scoring rules for each CARE-X model and its corresponding benchmark.

The significance of score differentials (ΔS is given by the scoring rule of a CARE-X model minus the scoring rule of the corresponding CE model) is then evaluated with the Conditional Predictive Ability (CPA) test of Giacomini and White (2006).

The CPA test can be implemented either as an unconditional test or as a conditional test of predictive ability. In the first case, a rejection of the null hypothesis, coupled with a negative score differential, suggests that the CARE-X model is on average more accurate than the CE model. In the second case, some explanatory variables can be included in the test equation to obtain a conditional test of predictive ability. Since traditional fuels, biofuels and agricultural markets are intertwined, we use a dummy variable based on net oil price increases (NOPI) to check if there are asymmetries in the forecasting performance of CARE-X models across quiet and turbulent phases of the energy markets.[7] A rejection of the null hypothesis provides evidence that the score differential is statistically different from zero. Hence, according to the sign of the differential, it is possible to determine whether, conditionally on the state of the oil market, CARE-X forecasts are more accurate than CE forecasts.

5. Empirical Results

Our empirical findings can be summarized as follows:

- In-sample results: returns on ethanol are Granger caused by returns on corn, soybeans and wheat. These results hold for the left tail and the centre, but not for the right tail of the distribution. There is evidence of a feedback relation between returns on soybeans and ethanol, but only for the centre of the distribution. Granger causality also runs from returns on ethanol to returns on soybeans, but this relation is limited to the right tail of the distribution. Lastly, Granger causality runs from returns on ethanol to returns on the price indices, but only in the right tail of the distribution.
- Out-of-sample evaluation: returns on field crops help forecasting the left tail and the centre of the distribution of ethanol returns, but not vice versa. The in-sample

[7] Net oil price increases proxy oil shocks. Following Hamilton (1996), we construct a dummy variable, $NOPI_t$, where $NOPI_t$=1 if the spot price of WTI crude oil in month t is higher than the maximum price recorded during the previous three years, and $NOPI_t$=0 otherwise. We include the lagged value of NOPI in the CPA test.

evidence of Granger causality running from ethanol to soybeans and to price indices is therefore not confirmed out-of-sample.

- No linkages between returns on ethanol and cattle have emerged.

Compared with previous analyses, the contribution of this study to the empirical literature on biofuels is novel, since it presents not only in-sample evidence, but also focuses on out-of-sample evaluation of density forecasts of biofuels and food prices. As just highlighted, this helps identifying some differences between in-sample and out-of-sample results.

The availability of reliable forecasts is crucial both for financial decision-making and for policy design. For instance, Chang et al. (2011) have shown how to construct an optimal biofuel dynamic portfolio that consists of crude oil, corn and soybeans that petroleum companies can use to handle variations in crude oil and energy-crops prices and volatilities. The ability of our models to predict the left tail of the distribution of returns on the price of ethanol could be exploited in these settings, in that large negative returns often initiate a period of increased volatility.

The focus on the performance of models out-of-sample is important also for designing effective policies. In their 2011 Agricultural Outlook the OECD and FAO Secretariats discuss several measures to reduce price volatility in agricultural markets. They suggest that one of the key ingredients to formulate policy responses to extreme volatility in agricultural markets is to enhance market transparency by improving information and surveillance systems on market prospects (see OECD/FAO, 2011). The enhancement of global monitoring systems, such as the FAO Global Information and Early Warning System clearly depends on the availability of good forecasts.

5.1 In-sample results

Estimates of CARE-X models (1a)-(2a) for the period January 1987 through March 2012 (December 2010 for variables CAT and PI2) are shown in Table 2.

[Table 2 about here]

From Panel (a), the coefficients associated to lagged returns on ethanol are in the vast majority of cases statistically insignificant, irrespective of the quantile. Moreover, neither the magnitude, nor the sign of the coefficients display any clear pattern across quantiles. For both price indices the coefficient on lagged ethanol returns is positive in the left tail of the distribution, then constantly decreases, becoming negative in the right tail. Looking at the CARE-X model for corn, lagged returns on ethanol have negative coefficients for quantiles below the median, while positive coefficients above the median. In the case of soybeans, we can see that price returns are negatively correlated with ethanol and that most of the coefficients in the right tail of the distribution are statistically significant only at the margin.

The Bonferroni statistic for testing the joint null hypothesis that all coefficients associated to lagged returns on ethanol are zero across quantiles is reported in the last row of Table 2 for each estimated model. The null hypothesis is rejected for both price indices only, due to the significance of the coefficients estimated within the extreme quantiles (0.85-0.95). Interestingly, it is not possible to reject the null hypothesis for soybeans, which is the only commodity showing some statistically significant coefficients associated to lagged returns on ethanol.

In summary, we are unable to find any empirical evidence of a relation between ethanol and field crops or cattle. Keeping in mind that those findings are against the presence of bivariate Granger causality running from ethanol to returns on price indices, corn, wheat and cattle, there is no reason to expect that ethanol can be fruitfully exploited to make predictions for these variables.

The lack of causality running from ethanol returns to field crops returns is consistent with the literature reviewed in Section 2. However, our results are more general, since we are allowed to conclude that returns on ethanol do not provide useful information for forecasting any part of the distribution of returns on field crops and cattle.

The pattern of coefficients reported in Panel (b) of Table 2 is completely different. In CARE-X models with returns on ethanol as dependent variables, lagged returns on price indices and field crops are statistically significant for most of the quantiles, while coefficients on cattle are always statistically insignificant. All coefficients are positive and tend to decrease and become statistically insignificant as one moves from the left to the right tail of the distribution. Price indices and the other field crops seem to have predictive power for the centre and the left tail of the distribution of the ethanol returns. On the contrary, none of the exogenous variables is statistically significant for large quantiles. Finally, the Bonferroni tests allow to reject the joint null hypothesis of absence of Granger causality for all variables, with the exception of cattle.

To conclude, there is no evidence of bivariate Granger causality running from ethanol to field crops and cattle. Conversely, field crops seem to Granger cause ethanol. More precisely, our results suggest that returns on field crops might be used to forecast the centre and the left tail of the distribution of returns on ethanol, with the exception of its right tail.

5.2 Density forecasts

Since each model (1a)-(1b) is estimated to match expectiles with quantiles α = 0.05, 0.10, 0.15, …, 0.90, 0.95 (i.e. 19 quantiles), 228 series of one-period ahead forecasts are computed. The size of the estimation sample is T_0 = 123, which corresponds to 40% of the total number of observations, T, for the returns on ethanol. For each model and quantile, a vector of H forecasts is obtained with a rolling window procedure. We start by estimating each model using observations from t = 1 to $t = T_0$, and calculate forecasts in $t = T_0 + 1$. Then, observations from t = 2 to $t = T_0 + 1$ are used to estimate each model and compute the corresponding forecasts in $t = T_0 + 2$. This algorithm is iterated until forecasts in $t = T$ are calculated. Due to different sample sizes, the forecast evaluation period varies across commodities. Specifically, April 1997-March 2012 (H = 180) for PI1 and field crops, while April 1997-December 2010 (H = 165) for PI2 and cattle.

We use the weighted quantile score function and the CPA test to evaluate the predictive performance of each model on the whole distribution[8].

The weighted versions of the quantile scoring rules assign a higher score to specific parts of the distribution. For instance, if one is interested on the tails of the distribution, higher penalty can be attached to forecast errors that occur outside the inter-quartile range and a lower penalty to errors around the median.

The numerical values of the score differentials and results of the unconditional (UPA) and conditional predictive ability (CPA) tests are presented in Table 3.

[Table 3 about here]

The score differentials reported in Panel (a) are generally positive and the UPA and CPA tests lead to 8 rejections out of 90 comparisons. That is, ethanol has in general no predictive content for (any part of) the distributions of the returns on field crops, cattle and price indices. In Panel (b), the average un-weighted ("Uniform") standardized score differential is negative, meaning that density forecasts from the CARE-X model (1b) are to be preferred to the

[8] We have also analysed the accuracy of quantile forecasts. More precisely, for each quantile and model, we computed the asymmetric quadratic loss and implemented a forecasts encompassing test. Results based on these alternative approaches are presented in the Appendix and are consistent with those obtained from density forecasts.

density forecasts from the benchmark CE model (2b). However, the UPA test indicates that the score differential is not statistically different from zero. Table 3 shows two variants of the CPA test. The first, CPA1, includes in the test function a constant term and the first lag of the score differential, while the second test, CPA2, is conditional on both a constant term and the NOPI dummy variable. The conditional tests always reject the null hypothesis of equal predictive ability, with p-values lower than 5% (6 out of 12 cases) or lower than 10% (6 out of 12 cases). These results suggest that corn can be used to forecast the whole distribution of ethanol.

Although the numerical values of the weighted score differentials are almost invariantly negative (with the only exception of cattle), in most cases we cannot reject, neither unconditionally nor conditionally, the null hypothesis of equal predictive ability for the tails (taken jointly) and the right tail of the distribution. Conversely, the null hypothesis is rejected most of the times in the centre and left tail of the distribution. These findings are, on the one hand, not surprising, since different weights are applied to different parts of the distribution, while, on the other hand, they contribute to illustrate that conditioning on past forecasting performances (i.e. CPA1) or on the state of the oil market (i.e. CPA2) is relevant for the output of the predictive ability tests.

The ability of field crops to forecast the centre and the left tail of the distribution of ethanol returns is confirmed as well. In particular, for PI1 and field crops, 35 rejections of the null hypothesis of equal predictive ability are observed out of 60 cases. The majority of those rejections is observed in the centre (18.3%) and in the left tail (20%). The predictive ability of PI2 and cattle is extremely poor. Actually, the null hypothesis of equal predictive ability is rejected 12 times out of 30 comparisons, most of which are due to PI2.

Figure 2 reports different score functions for the density forecasts of ethanol, obtained with CARE-X model (1b), where the explanatory variable is corn, and compared with the score of the density forecasts from the corresponding CE benchmark model (2b).

[Figure 2 about here]

The top Panel ("Uniform") displays the unweighted quantile score function. In this case forecast errors are given the same penalty along the whole support of the distribution. The score of the CE benchmark model (2b) lies above the score of the CARE-X model (1b) for all quantiles below $\alpha = 0.75$. As a consequence, the score differential is negative in the left tail and around the centre of the distribution, while positive in the right tail. This suggests, once again, that both the centre and the left tail of the distribution of ethanol returns can be predicted using corn. The four panels at the bottom of Figure 2 show weighted score functions. The two plots in the left part of the middle Panel illustrate quantile scores that assign more weight to forecast errors in the centre and in the left tail of the distribution. In this case, forecasts obtained from the CARE-X model are generally superior to forecasts calculated with the benchmark CE model. The two plots in the right part of the middle Panel show quantile scores which give more weight to forecast errors in both tails and in the right tail of the distribution. Focusing on the latter, the density forecasts obtained using the CE model (2b) are on average more accurate than the predictions generated by the corresponding CARE-X model (1b).

6. Conclusions

In this paper we have studied the relationship between returns on ethanol, field crops and cattle from the new perspectives of out-of-sample Granger causality and predictability. Instead of focusing on specific moments, we have analysed the whole distribution of returns both in-sample and out-of-sample.

This new line of investigation is appropriate for at least three reasons. First, while previous studies on the biofuels-food price relation carry out only in-sample Granger causality tests, our analysis correctly interprets the definition of causality provided by Granger (1969). More specifically, the definition of causality is a statement about forecasting ability, hence tests whose direct focus is on forecasting are more appropriate (Ashley et al., 1980) and often have more power than in-sample tests (Chen, 2005; Clark and McCracken, 2005).

Second, since returns are generally non-normal, a failure to reject the null hypothesis of absence of Granger causality in mean does not necessarily exclude the presence of Granger causality for higher moments of the distribution. Therefore, our contribution, which explicitly analyses the whole distribution of returns, can provide useful information for a variety of forecast users and forecast purposes.

Third, our work partly fills some gaps in the biofuel literature recently highlighted in the survey by Serra and Zilberman (2013). According to these authors, while most studies dealing with the biofuel-food price transmission have focused on price levels, price volatility has received little attention. Moreover, little is known about the transmission mechanism of biofuel price shocks to food price volatility and about the response of biofuel volatility to oil price shocks. Lastly, they point out that since biofuel feedstocks are used as production factors also for food products, the literature should consider not only the biofuels market chain, but also the transmission of price shocks along the food market chain.

Our study analyses predictability relationships between ethanol price returns and food price returns relying on density forecasts. Hence, it contributes to a better understanding of the links between first, second and higher moments of the distributions of ethanol and food price returns, under both normal and extreme market conditions (i.e. we use net oil price increases to assess if the forecasting performance of models varies through different phases of the oil market). Moreover, in order to extend the range of food commodities included in the analysis, we also considered cattle.

Quantile and density forecasts have been used to address the following questions: *i*) what is the direction of Granger causality between ethanol, field crops, and cattle? *ii*) are the observed causality linkages a feature of the whole return distribution or of some specific parts? *iii*) can any in-sample evidence of causality be exploited to improve out-of-sample forecasts?

Both in-sample and out-of-sample results confirm and extend the findings of most of the existing literature. Actually, we find very limited empirical evidence, if not any at all, that ethanol returns Granger cause field crops or cattle returns. Rather, we provide empirical support for the existence of reverse causality, running from field crops to ethanol.

More precisely, we show that ethanol returns cannot be used to forecast any part of the distributions of returns on field crops and cattle. Second, both quantile and density forecasts for ethanol can be improved by using returns on field crops as explanatory variables. Third, these results hold for the centre and the left tail of the distribution of ethanol returns, but not for its right tail. This last finding suggests that the information content of returns on field crops can be fruitfully exploited to forecast extreme ethanol price decreases, which are of interest in risk management to compute value-at-risk or expected shortfall. Evidence of predictability for the centre of the distribution is useful for constructing prediction intervals for policy evaluation exercises.

Lastly, it is worth recalling that forecasts have been evaluated on a sample running from the late 1997 through the early 2012. This time period is very challenging, since it is characterized by two recessions, very volatile energy markets and includes the financial crisis. Clearly, these events are mirrored in the recent developments of the whole U.S. economy, including the biofuel and agricultural markets and the price dynamics of the commodities traded in those markets. Any change that these markets might have experienced can hardly be identified by a single moment of the distribution of returns. Rather, it is more likely that, given the complexity and the magnitude of events such as recessions and financial crises, a clearer picture of the predictive relationships between ethanol, field crops and cattle returns can be obtained either by looking at their entire distributions, as we did in this paper, or with more flexible modelling approaches such as the wavelet coherence analysis of Vacha et al. (2013).

An interesting topic for future research would be to raise the bar of forecast evaluation by using these methods and other econometric models common in the biofuels literature in financial applications such as hedging, portfolio allocation or value-at-risk.

References

Ashley, R., Granger, C.W.J. and Schmalensee, R. (1980), "Advertising and Aggregate Consumption: An Analysis of Causality", *Econometrica*, 48, 1149-1167

Balcombe, K. (2011). "The nature and determinants of volatility in agricultural prices: an empirical study". In: Prakash, A. (Ed.), Safeguarding Food Security in Volatile Global Markets. FAO, Rome, pp. 85-106.

Bastianin, A., M. Galeotti and Manera, M. (2013) "Biofuels and food prices: searching for the causal link", USAEE Working Paper No. 13-120.

Cenesizoglu, T. and Timmermann, A. (2008), "Is the Distribution of Stock Returns Predictable?", Working Paper, University of California at San Diego

Chang, T.-H., H.-M. Su and Chiu, C.-L. (2011), "Value-at-risk estimation with the optimal dynamic biofuel portfolio", *Energy Economics*, 33, 264-272.

Chen, S.S. (2005), "A note on in-sample and out-of-sample tests for Granger causality", *Journal of Forecasting*, 24, 453-464.

Clark, T.E. and McCracken, M.W. (2005), "The power of tests of predictive ability in the presence of structural breaks", *Journal of Econometrics*, 124, 1-31.

Efron, B. (1991), "Regression Percentiles Using Asymmetric Least Squared Error Loss", *Statistica Sinica* 1, 93-125

Engle, R.F. and Manganelli, S. (2004), "CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles", *Journal of Business & Economic Statistics*, 22, 367-381

Esmaeili, A. and Shokoohi, Z. (2010). "Assessing the effect of oil price on world food prices: application of principal component analysis", *Energy Policy*, 39, 1022-1025.

Giacomini, R. and White, H. (2006), "Test of Conditional Predictive Ability", *Econometrica* 74, 1545-1578

Gilbert, C. L. (2010), "How to understand high food prices", *Journal of Agricultural Economics*, 61, 398-425

Gneiting, T. and Ranjan, R. (2011), "Comparing Density Forecasts Using Threshold- and Quantile-Weighed Scoring Rules", *Journal of Business and Economic Statistics* 29, 411-422

Granger, C.W.J. (1969), "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods", *Econometrica* 37, 424-438

Granger, C.W.J. and Pesaran, H.M. (2000), "Economic and Statistical Measures of Forecast Accuracy", *Journal of Forecasting* 19, 537-560

Granger, C.W.J. and Sin, C.-Y. (2000), "Modelling the Absolute Returns of Different Stock Indices: Exploring the Forecastability of an Alternative Measure of Risk", *Journal of Forecasting* 19, 166-187

Hamilton, J.D. (1996), "This is What Happened to the Oil Price-Macroeconomy Relationship", *Journal of Monetary Economics* 38, 215-220

Janda, K., Kristoufek, L. and Zilberman, D. (2012), "Biofuels: policies and impacts", *Agricultural Economics*, 58, 367-371.

Kretschmer, B. and Peterson, S. (2010) "Integrating bioenergy into computable general equilibrium models - A survey", *Energy Economics*, 32, 673-686.

Kristoufek, L., K. Janda, and Zilberman, D. (2012a), “Correlations Between Biofuels and Related Commodities Before and During the Food Crisis: A Taxonomy Perspective”, *Energy Economics*, 34, 1380-1391

Kristoufek, L., K. Janda and Zilberman, Z. (2012b), “Mutual Responsiveness of Biofuels, Fuels and Food Prices”, Working Paper 38, Centre for Applied Macroeconomic Analysis, Australian National University.

Kristoufek, L., K. Janda, and Zilberman, D. (2013a), “Non-linear Price Transmission between Biofuels, Fuels and Food Commodities”, Working Paper 481, Centre for Economic Research and Graduate Education - Economic Institute, Prague.

Kristoufek, L., Janda, K., and Zilberman, D. (2013b), “Regime-dependent topological properties of biofuels networks”, *European Physical Journal B*, 86(2), Article 40.

Kuan, C.-M., Yeh, J.H. and Hsu, Y.-C. (2009), “Assessing Value at Risk with CARE, the Conditional Autoregressive Expectile Models”, *Journal of Econometrics* 150, 261-270

Mercer-Blackman,V., Samiei, H. and Cheng, K. (2008), “Biofuel Demand Pushes Up Food Prices”, International Monetary Fund Survey Magazine: IMF Research, 17 October

Mitchell, D. (2008), “A Note on Rising Food Prices”, World Bank Policy Research Working Paper No. 4682, July

Nazlioglu, S. (2011), “World oil and agricultural commodity prices: evidence from nonlinear causality”, *Energy Policy*, 39, 2935-2943.

Nazlioglu, S., Erdem, C. and Soytas, U. (2013), “Volatility spillover between oil and agricultural commodity markets”, *Energy Economics*, 36, 658-666.

Nazlioglu, S. and Soytas, U. (2011). “Oil Price, agricultural commodity prices, and the dollar: a panel cointegration and causality analysis”, *Energy Economics*, 34, 1098-1104.

Newey, W.K. and Powell, J.L. (1987), “Asymmetric Least Squares Estimation and Testing”, *Econometrica* 55, 819-847

OECD/FAO (2011), “OECD-FAO Agricultural Outlook 2011-2020”, OECD Publishing and FAO. http://dx.doi.org/10.1787/agr_outlook-2011-en.

Parker, T. (2013), “Ethanol: One of Several Factors Fueling Corn Prices”, Yahoo! Finance, January, 29. Available at: http://finance.yahoo.com/news/ethanol-one-several-factors-fueling-194943041.html

Pesaran, H.M. and Skouras, S. (2002), “Decision-Based Methods for Forecast Evaluation”, in: Clements, M.P. and Hendry, D.F. (eds.), *A Companion to Economic Forecasting*. Blackwell: Oxford

Phillips, P.C.B. and Yu, J. (2011), “Dating the Timeline of Financial Bubbles During the Subprime Crisis”, *Quantitative Economics* 2, 455-491

Saghaian, S.H. (2010), “The Impact of the Oil Sector on Commodity Prices: Correlation or Causation?”, *Journal of Agricultural and Applied Economics* 42, 477-485

Serra, T. and Zilberman, D. (2013), “Biofuel-related price transmission literature: a review”, *Energy Economics*, 37, 141-151.

Serra, T., Zilberman, D., Gil, J.M. and Goodwin, B.K. (2011), “Nonlinearities in the U.S. Corn-Ethanol-Oil-Gasoline Price System”, *Agricultural Economics* 42, 35-45

Taylor, J.W. (2008), "Estimating Value at Risk and Expected Shortfall Using Expectiles", *Journal of Financial Econometrics* 6, 231-252

Ubilava, D. and Holt, M. (2010), "Forecasting Corn Prices in the Ethanol Era", mimeo, Department of Agricultural Economics, Purdue University

UNCTAD, United Nations Conference on Trade and Development (2008), "Addressing the Global Food Crisis: Key Trade, Investment and Commodity Policies in Ensuring Sustainable Food Security and Alleviating Poverty". Available at: http://www.unctad.org/en/docs/osg20081_en.pdf

Vacha, L., Janda, K., Kristoufek, L., and Zilberman, D. (2013), "Time-frequency dynamics of biofuel-fuel-food system", *Energy Economics*, 40, 233-241.

Wixson, S. E. and Katchova, A.L. (2012), "Price Asymmetric Relationships in Commodity and Energy Markets", mimeo. Paper presented at the 123rd European Association of Agricultural Economists, Dublin, Ireland

Zhang, Z., Lohr, L., Escalante, C. and Wetzstein, M. (2010), "Food Versus Fuel: What Do Prices Tell Us?", *Energy Policy* 38, 445-451

Zhang, Z., Lohr, L., Escalante, C. and Wetzstein, M. (2009), "Ethanol, Corn, and Soybean Price Relations in a Volatile Vehicle-Fuels Market", *Energies* 2, 320-339

Zhang, Z., Vedenov, D. and Wetzstein, M. (2007), "Can the U.S. Ethanol Industry Compete in the Alternative Fuels Market?", *Agricultural Economics* 37, 105-112

Zilberman, D., Hochman, G., Rajagopal, D., Sexton, S. and Timilsina, G. (2013), "The Impact of Biofuels on Commodity Food Prices: Assessment of Findings", *American Journal of Agricultural Economics*, 95, 275-281.

Tables & Figures

Table 1. Descriptive Statistics: January 1987 - December 2010/March 2012

Panel (a): Nominal Prices	ETH	PI1 (CAT excl.)	PI2 (CAT incl.)	COR	SOY	WHE	CAT
Mean	1.53	102.63	347.25	2.70	6.71	3.84	76.52
Coef. Var.	0.35	0.51	0.28	0.40	0.33	0.38	0.15
Min	0.89	47.46	215.08	1.43	4.00	1.99	58.60
Date Min	01/1987	01/1987	01/1987	02/1987	10/2001	11/1999	09/1998
Max	3.58	287.70	723.11	6.93	13.30	9.84	104.00
Date Max	06/2006	06/2011	12/2010	08/2011	08/2008	03/2008	12/2010
Panel (b): Returns							
	ETH	PI1 (CAT excl.)	PI2 (CAT incl.)	COR	SOY	WHE	CAT
Mean	0.09	0.15	0.06	0.09	0.05	0.08	0.19
Coef. Var.	82.05	38.80	55.55	60.25	93.56	66.21	17.67
Skewness	0.40	0.53	0.37	-0.62	-0.23	-0.54	0.02
Kurtosis	4.26	10.76	5.79	6.60	4.81	6.57	4.51

Notes: CAT = cattle; COR = corn; ETH = ethanol; PI1 = price index 1; PI2 = price index 2; SOY = soybean; WHE = wheat. Entries are descriptive statistics for nominal prices (Panel a) and percentage log returns (Panel b). Panel (a) shows the sample average (Mean), the coefficient of variation (Coef. Var.), the minimum (Min) and maximum (Max) price and their dates (Date Min and Date Max). Panel (b) shows the sample average (Mean), the coefficient of variation (Coef. Var.), the sample Skewness and Kurtosis for log-returns. The time period spanned by the monthly nominal spot price of CAT and PI2 is January 1987-December 2010, while the monthly nominal spot prices of COR, ETH, SOY, WHE and PI1 are observed from January 1987 to March 2012.

Table 2. Coefficient Estimates: March 1987 - December 2010/March 2012

Panel (a) Does ethanol Granger cause variable *j*?

			j			
α	PI1	PI2	COR	SOY	WHE	CAT
0.05	0.156	0.022	-0.046	0.001	0.041	-0.036
0.10	0.073	0.022	-0.074	-0.047	0.000	-0.015
0.15	0.052	0.018	-0.065	-0.054	0.004	-0.004
0.20	0.036	0.016	-0.053	-0.054	0.005	-0.002
0.25	0.028	0.014	-0.044	-0.053	0.013	0.000
0.30	0.016	0.010	-0.040	-0.051	0.018	0.001
0.35	0.007	0.008	-0.034	-0.052	0.023	0.002
0.40	-0.001	0.005	-0.029	-0.053	0.026	0.004
0.45	-0.006	0.002	-0.025	-0.054	0.028	0.007
0.50	-0.011	0.001	-0.020	-0.056	0.030	0.008
0.55	-0.018	-0.001	-0.016	-0.058*	0.029	0.011
0.60	-0.029	-0.003	-0.011	-0.061*	0.029	0.014
0.65	-0.037	-0.006	-0.005	-0.063*	0.029	0.016
0.70	-0.048	-0.013	0.001	-0.065*	0.027	0.022
0.75	-0.061	-0.018	0.005	-0.067*	0.027	0.026
0.80	-0.086	-0.026	0.012	-0.071*	0.024	0.034
0.85	-0.113*	-0.039	0.026	-0.081*	0.014	0.041
0.90	-0.174***	-0.072*	0.029	-0.095**	0.010	0.049
0.95	-0.362***	-0.168***	0.001	-0.119*	-0.016	0.087**
Bonferroni	0.002***	0.015**	1.000	0.471	1.000	0.409

Panel (b) Does variable *i* Granger cause ethanol?

			i			
α	PI1	PI2	COR	SOY	WHE	CAT
0.05	0.280***	0.422***	0.212**	0.333***	0.313***	0.224
0.10	0.234***	0.388***	0.248***	0.313***	0.273***	0.157
0.15	0.199***	0.311***	0.250***	0.316***	0.229***	0.082
0.20	0.177***	0.253***	0.244***	0.299***	0.178**	0.060
0.25	0.158***	0.224***	0.242***	0.285***	0.157**	0.051
0.30	0.144***	0.187**	0.238***	0.273***	0.144**	0.039
0.35	0.137**	0.175*	0.234***	0.255***	0.138**	0.034
0.40	0.132**	0.167	0.229***	0.245***	0.135**	0.030
0.45	0.127**	0.160	0.225**	0.231**	0.132**	0.027
0.50	0.122*	0.154	0.219**	0.222**	0.131*	0.024
0.55	0.118	0.150	0.214*	0.213*	0.130	0.016
0.60	0.114*	0.148	0.210**	0.205**	0.131**	0.012
0.65	0.112**	0.148	0.201**	0.197**	0.132**	0.010
0.70	0.110**	0.149	0.193**	0.188**	0.131**	0.015
0.75	0.108*	0.151	0.187**	0.183**	0.131**	0.024
0.80	0.105*	0.152	0.178**	0.178**	0.134**	0.024
0.85	0.104	0.153	0.166	0.157	0.138*	0.008
0.90	0.114	0.170	0.156	0.138	0.141	-0.031
0.95	0.123	0.073	0.105	0.060	0.167	-0.156
Bonferroni	0.000***	0.000***	0.001***	0.006***	0.002***	1.000

Notes: See notes to Table 1. In Panel (a) the estimated model is (1a), where the dependent variable is the returns on variable *j*. In Panel (b) the estimated model is (1b), where the dependent variable is the returns on ethanol. Entries are coefficient estimates (with the exception of the rows headed "Bonferroni", where entries are *p*-values). Stars denote rejection of the null hypothesis of no Granger Causality (GC) running from ethanol to variable *j* (Panel a) and from variable *i* to ethanol (Panel b). The null hypothesis is H_0: $\beta_1 = 0$ in model (1a) for Panel (a), or H_0: $\gamma_1 = 0$ in model (1b) for Panel (b). Headers reported in column α indicate the quantiles estimated from expectiles. "Bonferroni" indicates the Bonferroni bound for the joint null hypothesis of no GC across quantiles. * (**) [***] denotes rejection of the null hypothesis of no GC at 0.10 (0.05) [0.01] significance level.

Table 3. Density Forecasts: Score Function Differentials and CPA Tests

Panel (a) Does ethanol help forecasting variable *j*?

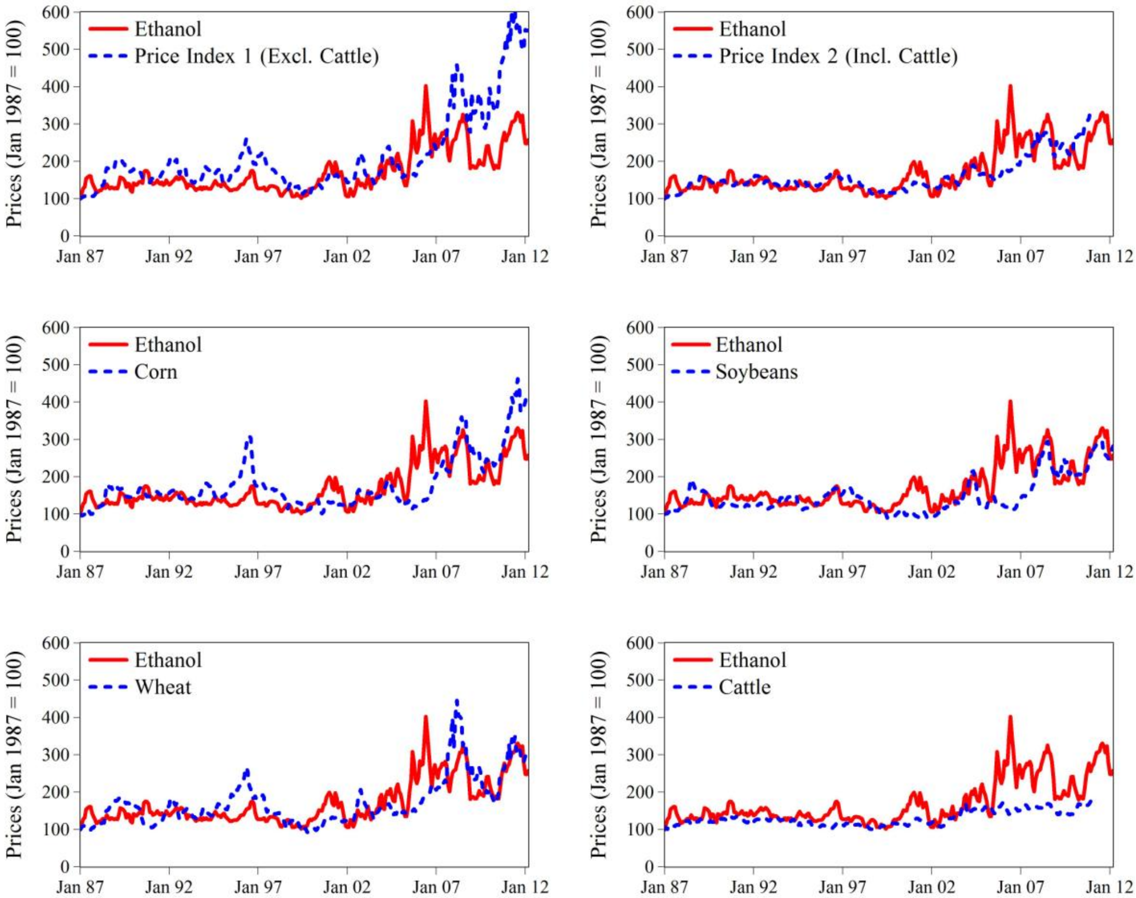
	Weights: Uniform			Weights: Center			Weights: Tails			Weights: Left Tail			Weights: Right Tail		
j	UPA	CPA1	CPA2	UPA	CPA1	CPA2	UPA	CPA1	CPA2	UPA	CPA1	CPA2	UPA	CPA1	CPA2
PI1	0.023	0.023	0.023	0.014	0.014	0.014	0.049	0.049	0.049	-0.032	-0.032	-0.032	0.080	0.080	0.080
PI2	-0.016	-0.016	-0.016	-0.025	-0.025	-0.025	0.014	0.014	0.014	-0.003	-0.003	-0.003*	-0.015	-0.015	-0.015
COR	0.043	0.044	0.043*	0.045	0.045	0.045*	0.034	0.034	0.034*	0.021	0.021	0.021	0.058	0.058	0.058*
SOY	-0.064	-0.064	-0.064	-0.063	-0.063	-0.063	-0.061	-0.061*	-0.061	-0.061	-0.061	-0.061	-0.056	-0.056	-0.056
WHE	0.099	0.099	0.099	0.092	0.092	0.092	0.108	0.108	0.108	0.056	0.056	0.056	0.132*	0.132**	0.132
CAT	-0.082	-0.082	-0.082	-0.086	-0.086	-0.086	-0.056	-0.056	-0.056	-0.105	-0.105	-0.105	-0.041	-0.041	-0.041

Panel (b) Does variable *j* help forecasting ethanol?

	Weights: Uniform			Weights: Center			Weights: Tails			Weights: Left Tail			Weights: Right Tail		
j	UPA	CPA1	CPA2	UPA	CPA1	CPA2	UPA	CPA1	CPA2	UPA	CPA1	CPA2	UPA	CPA1	CPA2
PI1	-0.135*	-0.135**	-0.135**	-0.147*	-0.147**	-0.147**	-0.095	-0.095	-0.095	-0.173**	-0.173**	-0.173***	-0.045	-0.045	-0.045
PI2	-0.097	-0.097*	-0.097*	-0.111	-0.111*	-0.111*	-0.055	-0.055	-0.055	-0.138*	-0.138	-0.138**	-0.016	-0.016*	-0.016
COR	-0.102	-0.102*	-0.102*	-0.111	-0.111**	-0.111**	-0.065	-0.065	-0.065	-0.147*	-0.147*	-0.147**	-0.018	-0.018*	-0.018
SOY	-0.163**	-0.163**	-0.163*	-0.179**	-0.179**	-0.179*	-0.107	-0.107	-0.107	-0.189**	-0.189**	-0.189**	-0.073	-0.073	-0.073
WHE	-0.143*	-0.143**	-0.143**	-0.155**	-0.155**	-0.155**	-0.103	-0.103	-0.103	-0.178**	-0.178**	-0.178***	-0.055	-0.055	-0.055
CAT	-0.064	-0.064*	-0.064**	-0.073	-0.073*	-0.073**	-0.036	-0.036	-0.036	-0.105	-0.105	-0.105**	0.003	0.003	0.003

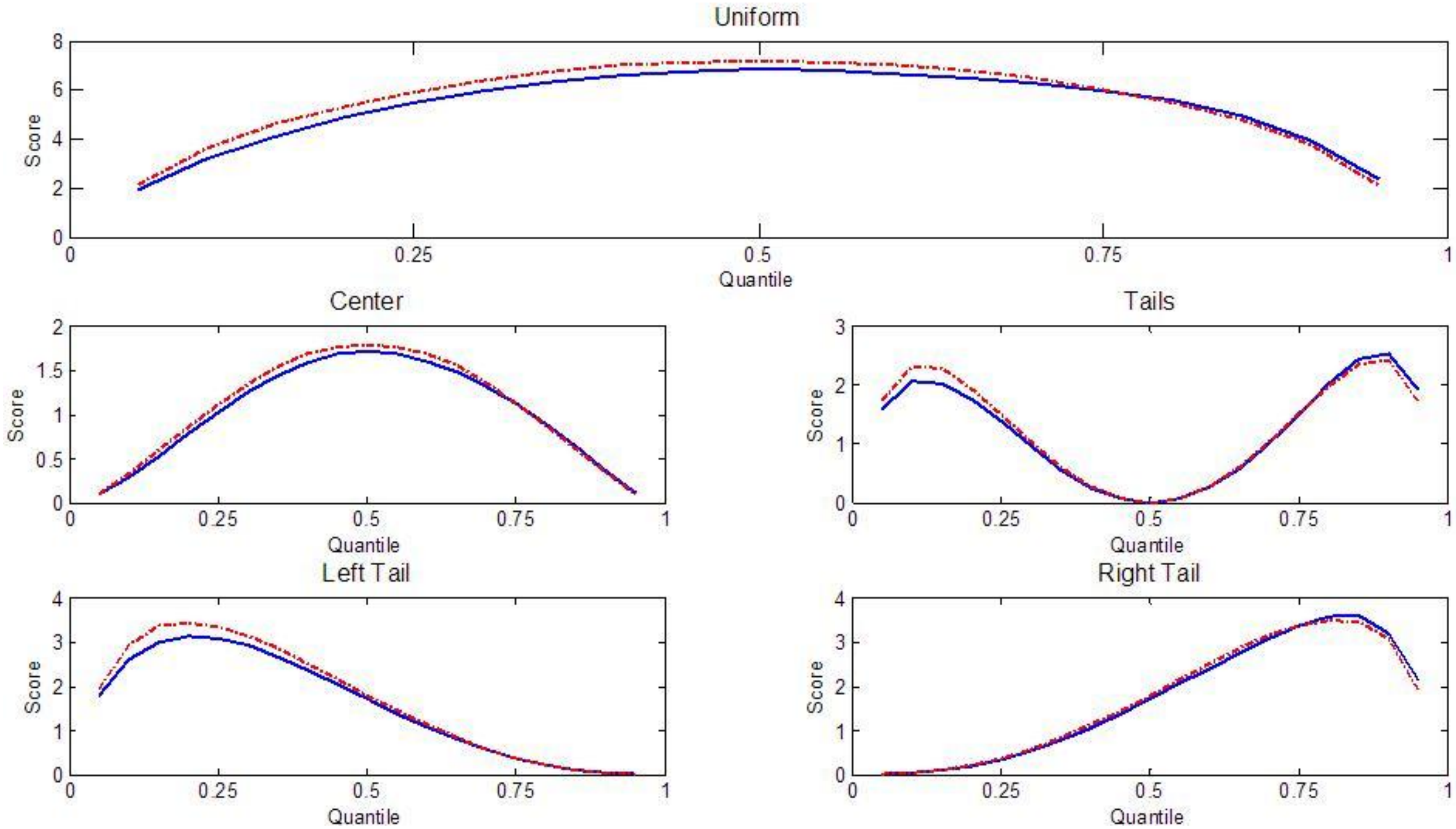
Notes: See notes to Table 1. This table reports standardized weighted quantile scoring rule differentials ΔS, (i.e. the scoring rule of a CARE-X model minus the scoring rule of the corresponding CE model). Weights are used to evaluate score differentials in the Centre, Tails (both), Left Tail and Right Tail of the distribution. The option "Weights: Uniform" considers the un-weighted scoring rule. In Panel (a) negative numbers (i.e. $\Delta S<0$) indicate that the scoring rule of CARE-X model (1a) is on average lower than the scoring rule of CE model (2a). In Panel (b) negative numbers (i.e. $\Delta S<0$) indicate that the score function of CARE-X model (1b) is on average lower than the score function of CE model (2b). Asterisks indicate rejection of the null hypothesis of the CPA test, namely H_0: $E(\Delta S)=0$. * (**) [***] denotes rejection of the null hypothesis at 10% (5%) [1%]. A rejection of the null, coupled with $\Delta S<0$, indicates that CARE-X forecast are on average more accurate than CE forecasts for a given part of the distribution of the dependent variable. UPA indicates the unconditional predictive ability test. CPA1 is the CPA test based on the lagged value of the score differential. CPA2 is the CPA test based on the lagged value of the Net Oil Price Increase (NOPI).

Figure 1. Prices: Ethanol, Indices, Field Crops and Cattle



Notes: Prices are represented on a common scale (i.e. nominal prices have been multiplied by 100 and divided by their value in January 1987).

Figure 2: Quantile Score Functions for Ethanol Forecasts



Notes: This figure shows average weighted quantile score functions for ethanol forecasts obtained with equations (1b) and (2b) and corn as the explanatory variable. A continuous line identifies scores associated to density forecasts from CARE-X model (1b), while a dash-dotted line is used for density forecast from CE benchmark model (2b). CARE-X forecasts are preferred to CE forecasts if CARE-X score lies below CE score.

Appendix to “Causality and Predictability in Distribution: the Ethanol-Food Price Relation Revisited”

Andrea Bastianin[*,**], Marzio Galeotti[***], Matteo Manera[*]

This appendix presents additional details and results for the paper “Causality and Predictability in Distribution: the Ethanol-Food Price Relation Revisited”.

[*] University of Milan-Bicocca and Fondazione Eni Enrico Mattei, Milan, Italy.
[**] Corresponding Author: Andrea Bastianin, DEMS - Department of Economics, Management and Statistics, University of Milan-Bicocca, Via Bicocca degli Arcimboldi, 8, I-20126, Milan, Italy. E-mail: andrea.bastianin@unimib.it. Phone: +39 02 64455879.
[***] University of Milan and IEFE-Bocconi, Milan, Italy.

A.1 Data description

Table A1. Data description

Series ID	Description	Unit	Frequency	Time Period	Source[a]
PE	Ethanol: Average Rack Prices F.O.B. Omaha, Nebraska	Dollars per Gallon	Monthly	Jan/1982 - Mar/2012	NEO
PC	Corn (Grain): Price Received	Dollars per Bushel	Monthly	Jan/1982 - Mar/2012	USDA
PS	Soybeans: Price Received	Dollars per Bushel	Monthly	Jan/1982 - Mar/2012	USDA
PW	Wheat: Price Received	Dollars per Bushel	Monthly	Jan/1982 - Mar/2012	USDA
PB	Cattle (>500 LBS): Price Received	Dollars per CWT[b]	Monthly	Jan/1982 - Dec/2010	USDA
YC[c]	Corn (Grain): Production	Dollars	Yearly	1982-2012	USDA
YS[c]	Soybeans: Production	Dollars	Yearly	1982-2012	USDA
YW[c]	Wheat: Production	Dollars	Yearly	1982-2012	USDA
YB[c]	Cattle (Incl Calves): Production	Dollars	Yearly	1988-2012	USDA

Notes: (a) NEO = Nebraska Energy Office; USDA = U.S. Department of Agriculture - National Agricultural Statistics Service; (b) CWT = hundredweight; (c) The value for 2012 is obtained as a cubic trend forecast.

The production variables described in Table A1 have been used to construct two commodity price indexes whose aim is to provide a summary of the price developments for field crops and cattle. The first (PI1 in the paper) is formed using percentage price variations of corn, wheat and soybeans; the second (PI2 in the paper) includes also cattle prices.

Both indices have been constructed by averaging prices with production based weights of the form:

$$w_{j,1,t} = \mathrm{Y}j_{,t} \,/\, (\mathrm{YC}_t + \mathrm{YW}_t + \mathrm{YS}_t), \qquad \text{for } j = \mathrm{C, S, W} \text{ and } t = 1982, \ldots, 2012 \qquad \text{(A1)}$$

$$w_{i,2,t} = \mathrm{Y}i_{,t} \,/\, (\mathrm{YC}_t + \mathrm{YW}_t + \mathrm{YS}_t + \mathrm{YB}_t), \qquad \text{for } i = \mathrm{C, S, W, B} \text{ and } t = 1982, \ldots, 2010 \qquad \text{(A2)}$$

Given that production variables are recorded at yearly frequency, we constructed monthly observations by assuming constant weights within the year (e.g. $w_{j,1,1/1982} = w_{j,1,2/1982} = \ldots = w_{j,1,12/1982}$, where r/1982 indicates the r-th month of year 1982).

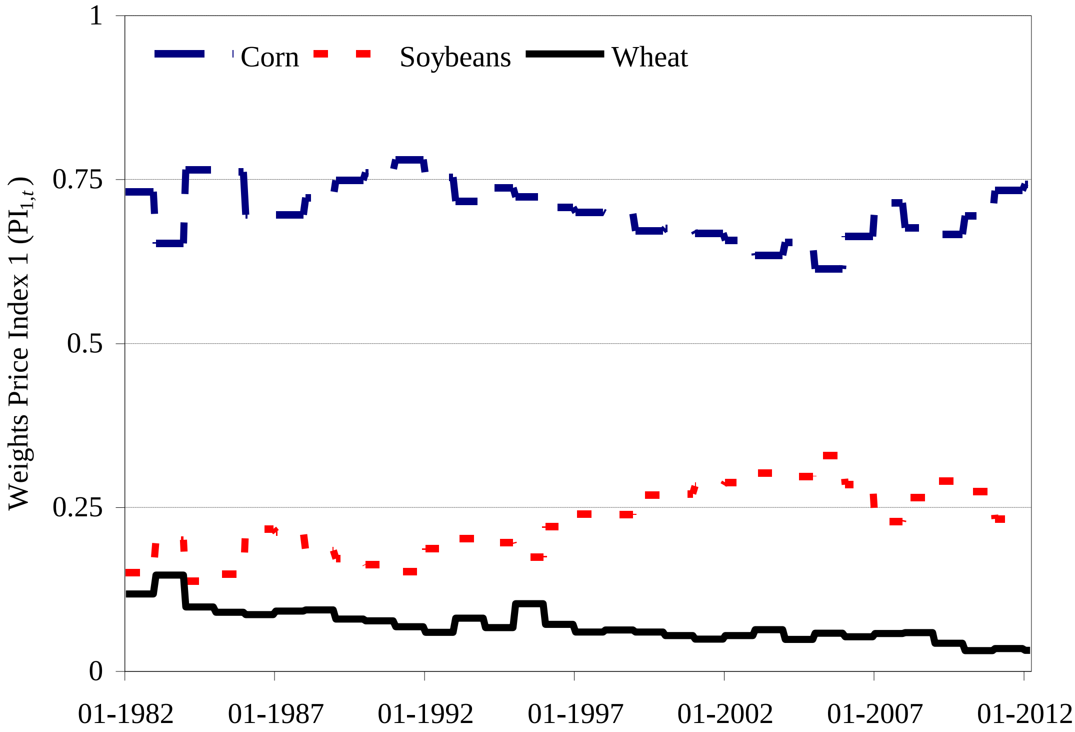
Weights calculated using current dollar production data are displayed in Figure A1. Current dollars price indices PI1 and PI2 are calculated as follows:

$$\mathrm{PI1}_{,t} = (\mathrm{PC}_t / w_{C,1,t}) + (\mathrm{PW}_t / w_{W,1,t}) + (\mathrm{PS}_t / w_{S,1,t}), \qquad \text{for } t = 1/1982, \ldots, 3/2012, \qquad \text{(A3)}$$

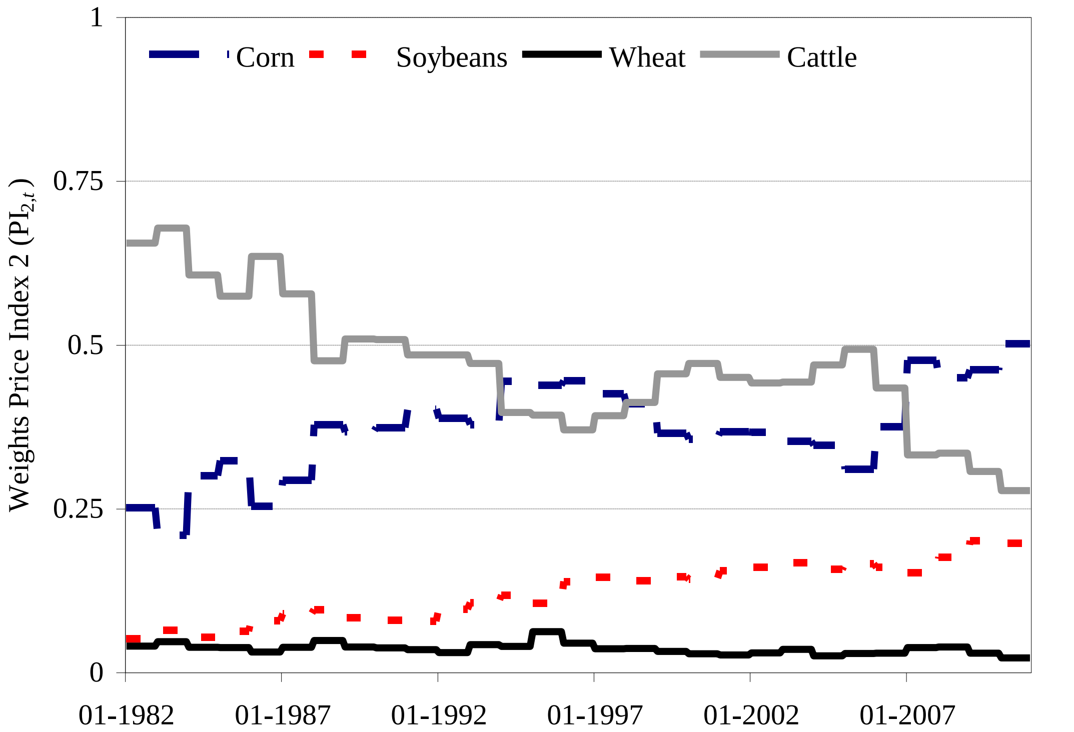
$$\mathrm{PI2}_{,t} = (\mathrm{PC}_t / w_{C,2,t}) + (\mathrm{PW}_t / w_{W,2,t}) + \ldots \qquad \text{for } t = 1/1982, \ldots, 12/2010. \qquad \text{(A4)}$$

$$\ldots + (\mathrm{PS}_t / w_{S,2,t}) + (\mathrm{PB}_t / w_{B,2,t})$$

Figure A1. Weights for Price Index 1 (Panel a) and Price Index 2 (Panel b).



Panel (a)



Panel (b)

A.2 Unit root tests

Table A2. Unit root and stationarity tests

Panel (a): log-prices						
Test:	ADF		Phillips-Perron		KPSS	
Exogenous:	C	C & T	C	C & T	C	C & T
Ethanol	0.2012	0.0473**	0.1696	0.0539*	NS	NS
Price Index 1 (Excl. Cattle)	0.8145	0.7809	0.7453	0.6441	NS	NS
Price Index 2 (Incl. Cattle)	0.9536	0.9508	0.9280	0.9085	NS	NS
Corn	0.5011	0.5270	0.7000	0.7327	NS	NS
Soybeans	0.4432	0.4570	0.5094	0.5425	NS	NS
Wheat	0.4503	0.5380	0.2591	0.2760	NS	NS
Cattle	0.7296	0.0821*	0.5386	0.3742	NS	NS

Panel (b): First difference of log-prices			
Test:	ADF	Phillips-Perron	KPSS
Exogenous:	N	N	C
Ethanol	0.0000***	0.0000***	S
Price Index 1 (Excl. Cattle)	0.0000***	0.0000***	S
Price Index 2 (Incl. Cattle)	0.0000***	0.0000***	S
Corn	0.0000***	0.0000***	S
Soybeans	0.0000***	0.0000***	S
Wheat	0.0000***	0.0000***	S
Cattle	0.0000***	0.0000***	S

Notes: Entries in the columns labelled “ADF” and “Phillips-Perron” are p-values of the null hypothesis that a series has a unit root. In the case of the KPSS test “NS” (“S”) denotes rejection (non rejection) of the null hypothesis of trend stationarity at 95% confidence level. “N” (neither trend, nor constant), “C” (constant) and “ C&T” (constant and trend) indicate the deterministic component included in the test equation.

To investigate the statistical properties of the log-price series Table A2 shows tests proposed by Dickey and Fuller (1979) in its augmented form (ADF), by Phillips and Perron (1988) (PP), and by Kwiatkowski, Phillips, Schmidt, and Shin (1992) (KPSS).

Panel a shows that all (log) prices have a unit root, while Panel b shows that their first difference is stationary.

A.3 Quantile forecasts

Since each model (1a-1b) (see Equations in the paper) is estimated to match expectiles with quantiles α = 0.05, 0.10, 0.15, …, 0.90, 0.95 (i.e. 19 quantiles), 228 series of one-period ahead forecasts are computed.[1] The size of the estimation sample is T_0 = 123, which corresponds to 40% of the total number of observations, T, for the returns on ethanol. For each model and quantile, a vector of H forecasts is obtained with a rolling window procedure. We start by estimating each model using observations from t = 1 to $t = T_0$, and calculate forecasts in $t = T_0 + 1$. Then, observations from t = 2 to $t = T_0 + 1$ are used to estimate each model and compute the corresponding forecasts in $t = T_0 + 2$. This algorithm is iterated until forecasts in $t = T$ are calculated. Due to different sample sizes, the forecast evaluation period varies across commodities. Specifically, April 1997-March 2012 (H = 180) for PI1 and field crops, while April 1997-December 2010 (H = 165) for PI2 and cattle.

For each quantile and model, we compute the asymmetric quadratic loss:

$$L_{t,\alpha}(r_{it}, \hat{q}_{\alpha,t}) = [\alpha + (1\text{-}2\alpha) \times I(r_{it} - \hat{q}_{\alpha,t} \leq 0)] \times |r_{it} - \hat{q}_{\alpha,t}|^2 \tag{A5}$$

where $\hat{q}_{\alpha,t}$ is the one-step ahead quantile forecast obtained from the ALS estimation of CARE-X models (1a)-(1b) and CE models (2a)-(2b); $t = T_0 + 1, \ldots, T$; α = 0.05, 0.10, 0.15, …, 0.90, 0.95; i = ETH, COR, SOY, WHE, CAT, PI1, PI2.

For any quantile α, a given CARE-X model produces more accurate forecasts than its benchmark if the average loss (defined as: $L_\alpha(.) = H^{-1} \sum^{T}_{t=T0+1} L_{t,\alpha}(.)$) of the CARE-X model is smaller than the average loss of the corresponding CE model.

This allows us to check whether the forecasts obtained with CARE-X models are on average more accurate than the benchmark CE forecasts.

In Panel (a) of Table A4 we show the asymmetric quadratic loss for CARE-X models (1a), where (lagged) ethanol returns is the exogenous explanatory variable and returns on variable j is the dependent variable. Since ethanol has no in-sample predictive power for both price indices, cattle and most field crops, we do not expect to find evidence of out-of-sample predictability running from ethanol to these commodities. When price indices PI1 and PI2 and field crops are considered, the benchmark CE model (2a) leads to an average loss lower than the average loss calculated on the CARE-X (1a) forecasts in 33 cases out of 50.

[1] This figure excludes the forecasts obtained using the benchmark models.

Table A4. Asymmetric Quadratic Loss

Panel (a) CARE-X model (1a) vs CE model (2a)										
j/α	0.05	0.10	0.25	0.40	0.50	0.60	0.75	0.80	0.90	0.95
PI1	11.147*	15.129*	22.697*	27.272	28.359	28.802	27.076	25.333	19.909	16.378
CE(1)	11.225	15.281	22.816	27.194*	28.129*	28.394*	26.338*	24.501*	18.804*	15.711*
PI2	3.346	4.832	7.676	9.004	9.391	9.475	8.759	8.218	6.429	4.682
CE(2)	3.270*	4.757*	7.454*	8.728*	9.036*	9.047*	8.226*	7.654*	5.859*	4.398*
COR	7.413*	10.227*	14.372	15.863	15.857	15.265	13.120	11.949	8.353*	6.001*
CE(1)	9.013	10.468	14.263*	15.698*	15.649*	15.093*	12.977*	11.812*	8.408	6.063
SOY	7.545*	10.014*	13.827*	15.126*	15.094*	14.424*	12.157*	11.006*	7.753*	5.308*
CE(1)	7.768	10.655	14.986	16.323	16.183	15.414	12.806	11.472	8.062	5.711
WHE	13.400	17.019	25.234	28.608	29.453	28.795	25.267	23.531	16.979	11.957
CE(1)	11.032*	15.400*	23.421*	26.618*	27.425*	26.810*	23.798*	22.162*	16.416*	11.714*
CAT	2.515*	3.509*	4.716*	5.141*	5.232*	5.149*	4.645*	4.294*	3.369*	2.590*
CE(2)	2.583	3.677	5.070	5.583	5.656	5.518	4.942	4.517	3.400	2.735
Panel (b) CARE-X model (1b) vs CE model (2b)										
j/α	0.05	0.10	0.25	0.40	0.50	0.60	0.75	0.80	0.90	0.95
PI1	13.286*	19.956	31.947	38.318	40.196	39.942	36.053	33.865	24.773	18.033
PI2	14.224	21.905	35.671	42.222	44.073	43.607	39.040	36.460	26.107	18.204
COR	13.754	20.544	33.330	39.237	40.890	40.322	36.602	34.162	25.318	18.371
SOY	13.404	19.545*	31.809*	37.852*	39.553*	39.452	35.752	33.505	25.113	18.084
WHE	13.823	20.186	32.293	38.525	40.275	40.068	36.014	33.567	25.092	17.854
CAT	14.673	22.108	36.020	42.490	44.196	43.752	39.076	36.185	25.378	17.966
CE(1)	15.915	24.804	37.023	42.991	44.052	42.985	37.284	34.453	24.414*	17.357
CE(2)	15.886	24.849	38.869	44.885	46.294	45.296	39.299	36.186	24.992	16.803*
EW-ALL	13.468	20.071	32.369	38.411	40.174	39.892	35.956	33.638	24.790	17.734
EW-CROPS	13.420	19.718	31.882	37.872	39.611	39.377*	35.627*	33.282*	24.867	17.748

Notes: See notes of Table 1. The table reports the asymmetric quadratic loss function for each estimated model. In Panel (a) the model of interest is (1a), where the dependent variable is the returns on variable *j*. In Panel (b) the model of interest is (1b), where the dependent variable is the returns on ethanol. Headers reported in row α indicate the quantiles estimated from expectiles. The benchmark forecasts are obtained from the CE models (2a) (Panel a) and (2b) (Panel b), and the equally-weighted forecast combinations (EW). Two are the evaluation periods: 1) April 1997-March 2012 (H = 180) for PI1 and field crops; 2) April 1997-December 2010 (H = 165) for PI2 and CAT. CE(1) and CE(2) refer to CE models evaluated in period 1) or period 2), respectively. EW-ALL and EW-CROPS are the EW combined forecasts based on all variables and field crops only. In Panel (a) an asterisk identifies the best model (i.e. lowest loss model) for each variable *j* and each quantile α. In Panel (b) an asterisk identifies the best model for each quantile α.

Ten of these occurrences are associated to the CARE-X models applied to soybeans returns, with respect to which our in-sample analysis has suggested that ethanol might have predictive power. CARE-X models (1a) applied to corn returns and PI1 produce more accurate forecasts that their corresponding benchmarks for extreme quantiles. However, in general the magnitude of the loss differentials is negligible.

For PI2 and wheat the CE benchmark (2a) is always associated to lower losses than the CARE-X (1a) models. Somewhat puzzling, the CARE-X model outperforms the benchmark also for cattle, with respect to which ethanol has no in-sample predictive power.

Panel (b) of Table A4 presents the asymmetric quadratic losses for ethanol forecasts, that is for CARE-X models (1b) and corresponding CE benchmarks (2b). Since field crops have no in-sample predictive power for the right tail of distribution of ethanol returns, we compute for each quantile two additional forecast models. The first (EW-ALL) is an equally weighted average of all forecasts for ethanol obtained from CARE-X models (1b), while the second (EW-CROPS) is an equally weighted average of forecasts for ethanol obtained from the subset of CARE-X models which include field crops as exogenous variables.

The benchmark CE models are outperformed 80% of the cases. CE models perform best only in correspondence to the 0.90 and 0.95 quantiles, confirming that extreme ethanol price increases cannot be predicted with field crops. The combined forecast model EW-CROPS performs best for some quantiles above the median (i.e. 0.60, 0.75, 0.80), where lack of in-sample Granger causality is found.

An alternative way of comparing the forecasting performance of CE and CARE-X models is to calculate optimal combining weights. For each model and quantile, the optimal combining weights are the estimated coefficients of regressing realized returns on the i-th variable on a constant term and forecasts obtained with CARE-X and CE models applied to the i-th variable, i = ETH, COR, SOY, WHE, CAT, PI1, PI2. As shown by Elliott and Timmermann (2004), when the loss function is asymmetric quadratic, the optimal forecast combination weights can be estimated with the Iterated Weighted Least Squares (IWLS) algorithm. If the optimal combining weight of the forecasts obtained with CARE-X model (1a) (ϕ_1) is equal to one and the optimal combining weight of the forecasts obtained with CE model (2a) (ϕ_2) is equal to zero, then the forecasts obtained with CARE-X model (1a) are more accurate than the forecasts based on the CE benchmark model (2a). In this sense, CARE-X model (1a) "forecast-encompasses" CE model (2a).

Table A5. Optimal Combining Weights: CARE-X and CE Forecasts

Panel (a)

	α													
	0.05		0.10		0.25		0.50		0.75		0.90		0.95	
j	CARE-X	CE	CARE-X	CE	CARE-X	CE	CARE-X	CE	CARE-X	CE	CARE-X	CE	CARE-X	CE
PI1	0.423*	-0.518	0.970	-1.887**	0.945	-1.886	0.295	2.299	-0.243**	2.287	-0.982***	2.678***	-0.790***	0.765
PI2	0.278**	-4.422***	0.018	-1.196	-0.003**	-6.054**	0.117	0.814	-0.039**	3.797**	-0.262**	4.745***	-0.030**	1.554**
COR	0.505**	0.732	0.459***	1.396	0.475***	0.607	0.431***	1.951	0.449***	1.852	0.428**	1.728	0.092***	0.599
SOY	0.246***	0.023	0.580	-1.292	0.740	-2.661	0.730	-0.024	0.699	-0.084	0.519*	0.216	0.290**	0.844
WHE	-0.706***	0.448	-0.634***	-0.149	-0.356***	-2.326	-0.173**	-0.678	-0.091***	0.761	0.203*	0.549	0.358	-0.072
CAT	0.577*	-0.064	0.766	-6.602	0.975	-3.581	0.938	-0.936	0.847	-1.685	0.457	-5.537***	0.374**	-2.475**

Panel (b)

	α													
	0.05		0.10		0.25		0.50		0.75		0.90		0.95	
j	CARE-X	CE	CARE-X	CE	CARE-X	CE	CARE-X	CE	CARE-X	CE	CARE-X	CE	CARE-X	CE
PI1	0.847	-0.142	0.930	-0.394	1.044	-0.142	1.016	0.133	0.770	0.406	0.293	0.660	0.161**	0.838
PI2	0.702	1.100*	0.865	0.402	0.981	-0.025	0.889	0.136	0.621	0.187	0.198	0.393	0.193*	0.279
COR	0.751	0.356	0.947	-0.310	0.900	-0.207	0.788	0.154	0.597	0.345	0.183*	0.618	0.053**	0.714
SOY	0.761	0.040	0.878	-0.102	0.978	-0.144	0.988	0.186	0.808	0.427	0.163*	0.604	0.036**	0.705
WHE	0.774	-0.063	0.966	-0.515	1.103	-0.196	1.038	0.088	0.790	0.455	0.270*	0.751	0.201**	0.830
CAT	0.655**	0.937	0.856	0.076	0.854	-0.107	0.753	0.057	0.604	0.069	0.270	0.340	0.177*	0.344

Notes: See notes to Table 1. In Panel (a) CARE-X and CE indicate models (1a) and (2a) respectively, where the dependent variable is the returns on variable *j*. In Panel (b) CARE-X and CE indicate models (1b) and (2b) respectively, where the dependent variable is the returns on ethanol. Headers reported under the label α indicate the quantiles estimated from expectiles. Numbers reported in Panel (a) are the combining weights ϕ_1 and ϕ_2 estimated from the regression model: $r_{jt} = \phi_0 + \phi_1 r_{jt}^{\text{CARE-X}} + \phi_2 r_{jt}^{\text{CE}} + e_{jt}$, where r_{jt} are actual returns from variable *j*, $r_{jt}^{\text{CARE-X}}$ are forecasts from CARE-X model (1a) and r_{jt}^{CE} are forecast from CE model (2a). If the single null hypotheses ϕ_1=1 and ϕ_2=0 are not rejected, then forecasting with CARE-X model (1a) is more accurate than forecasting with CE model (2a). In Panel (b) CARE-X and CE indicate models (1b) and (2b) respectively, where the dependent variable is the returns on ETH. Numbers reported in Panel (b) are the combining weights ψ_1 and ψ_2 estimated from the regression model: $r_{ETHt} = \psi_0 + \psi_1 r_{ETHt}^{\text{CARE-X}} + \psi_2 r_{ETHt}^{\text{CE}} + e_{ETHt}$, where r_{ETHt} are actual returns from ETH, $r_{ETHt}^{\text{CARE-X}}$ are forecasts from CARE-X model (1b) and r_{ETHt}^{CE} are forecast from CE model (2b). If the single null hypotheses ψ_1=1 and ψ_2=0 are not rejected, then forecasting with CARE-X model (1b) is more accurate than forecasting with CE model (2b). Coefficients ϕ_0, ϕ_1, ϕ_2, ψ_0, ψ_1 and ψ_2 are estimated with Iterated Weighted Least Squares. * (**) [***] denotes rejection of each single null hypothesis at 0.10 (0.05) [0.01] significance level.

Analogously, if the optimal combining weight of the forecasts obtained with CARE-X model (1b) (ψ_1) is equal to one and the optimal combining weight of the forecasts obtained with CE model (2b) (ψ_2) is equal to zero, then CARE-X model (1b) "forecast-encompasses" CE model (2b).

Estimated optimal combining weights and statistical tests of the null hypotheses ϕ_1=1 (ψ_1=1) and ϕ_2=0 (ψ_2=0) are reported in Table A5. Results in Panel (a) indicate that in most cases the CARE-X combining weights are statistically different from unity, suggesting lack of "forecast encompassing". Therefore, the test results are supportive of the in-sample absence of Granger causality running from ethanol to corn, wheat, price indices and cattle. Once again, ethanol seems to be useful to forecast some parts of the distribution of soybean returns. Actually, with the exception of the smallest and largest quantiles, the null hypotheses ϕ_1=1 and ϕ_2=0 are never rejected.

The results reported in Panel (b) have a penchant for "forecast encompassing". In the case of PI1 and field crops, the null hypotheses ψ_1=1 and ψ_2=0 cannot be rejected for quantiles from 0.05 to 0.75. Interestingly, for PI1 and PI2, which are by construction linear combinations of different series of returns, the two null hypotheses are not rejected also for the 0.90-th quantile.

A4. References

Dickey, D.A. and W.A. Fuller (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root", *Journal of the American Statistical Association* 74, 427-431

Elliott, G. and Timmermann, A. (2004), "Optimal Forecast Combinations under General Loss Functions and Forecast Error Distributions", *Journal of Econometrics* 122, 47-79

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt, and Y. Shin (1992), "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?", *Journal of Econometrics* 54, 159-178

Phillips, P.C.B. and P. Perron (1988), "Testing for a Unit Root in Time Series Regression", *Biometrika* 75, 335-346

Fig. 01 (Col)
Click here to download high resolution image

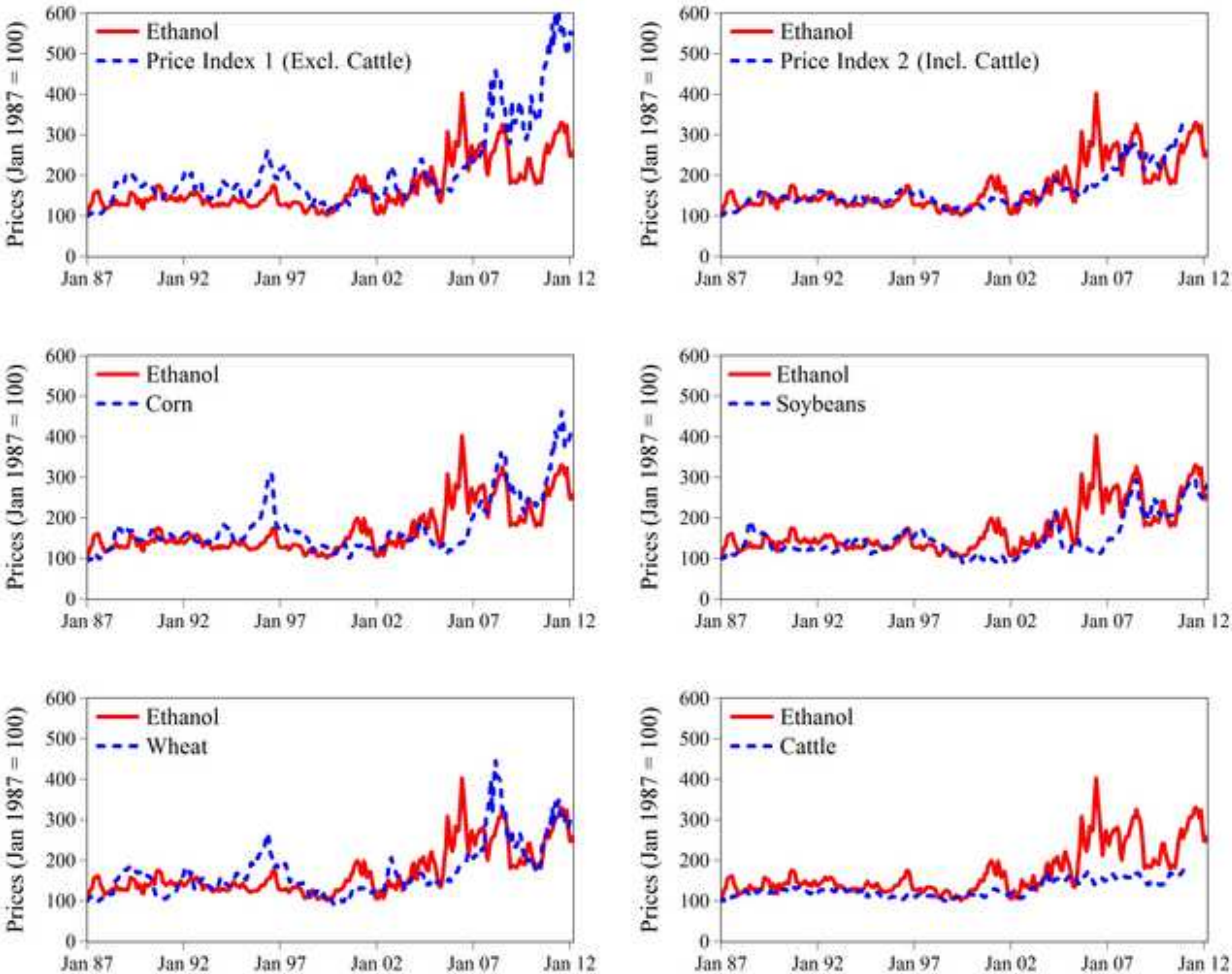


Fig. 02 (Col)
Click here to download high resolution image

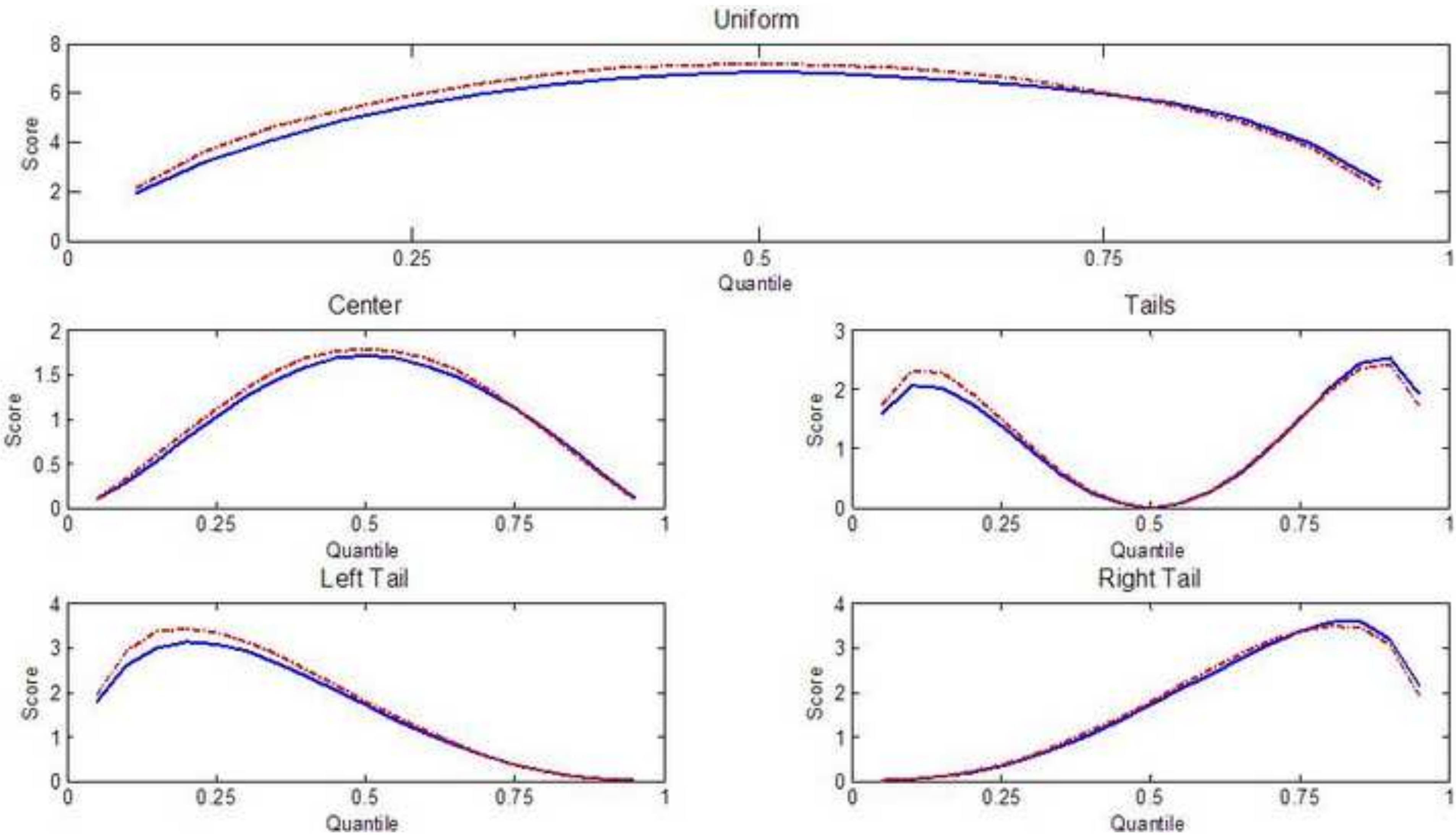


Fig. 02 (BW)
Click here to download high resolution image

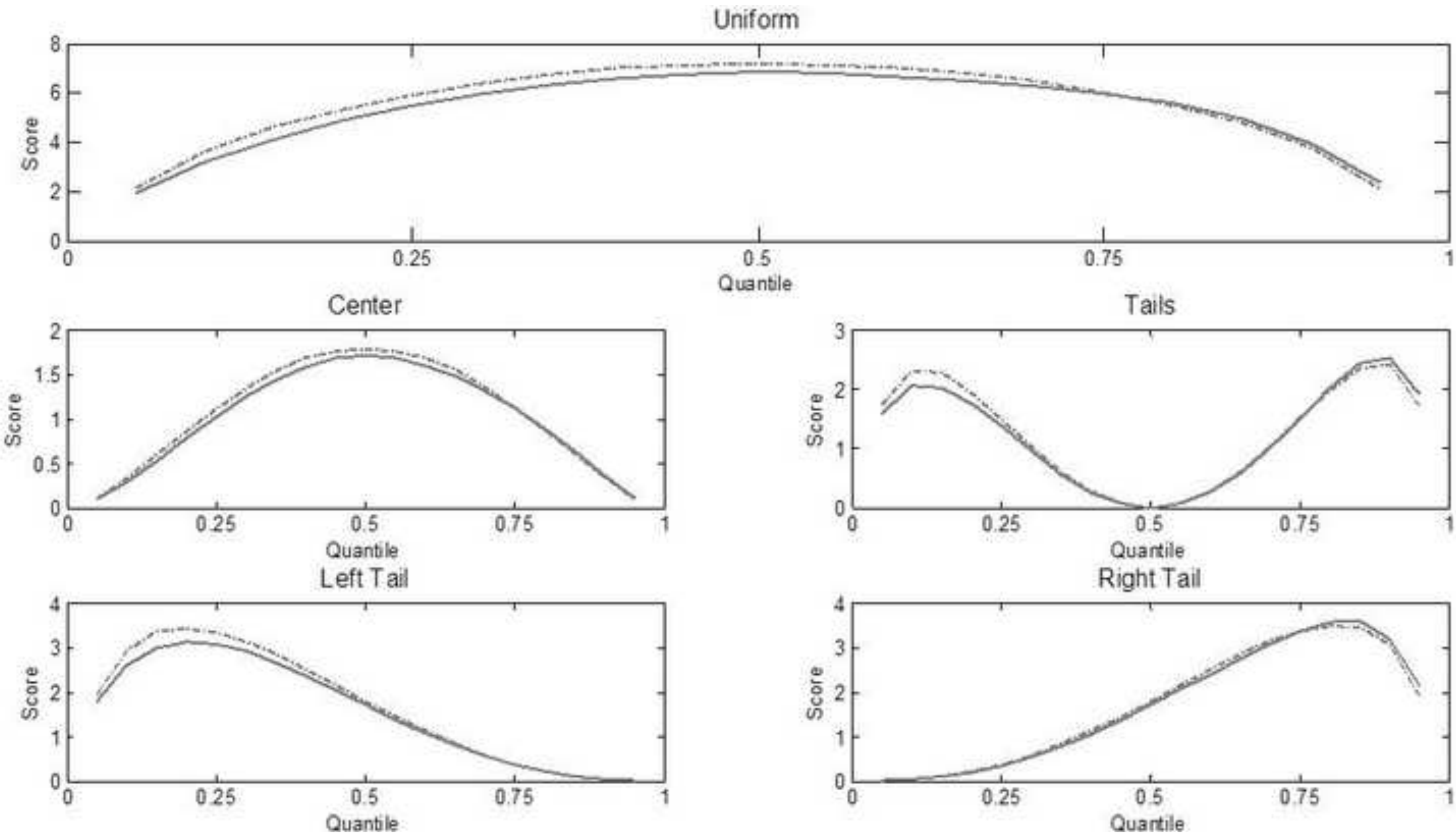


Fig. 01 (BW)
Click here to download high resolution image