



DEPARTMENT OF ECONOMICS,  
MANAGEMENT AND STATISTICS  
UNIVERSITY OF MILAN – BICOCCA

DEMS WORKING PAPER SERIES

**Food versus Fuel:  
Causality and Predictability in  
Distribution**

Andrea Bastianin, Marzio Galeotti, Matteo Manera

No. 241 – March 2013

Dipartimento di Economia, Metodi Quantitativi e Strategie di Impresa  
Università degli Studi di Milano - Bicocca  
<http://dems.unimib.it/>

# Food versus Fuel: Causality and Predictability in Distribution

Andrea Bastianin<sup>\*</sup>, Marzio Galeotti<sup>†</sup>, Matteo Manera<sup>\*‡</sup>

This Draft: March, 14<sup>th</sup> 2013

## Abstract

---

This paper examines the relationship between biofuels and commodity food prices in the U.S. from a new perspective. While a large body of literature has tried to explain the linkages between sample means and volatilities associated with ethanol and agricultural price returns, little is known about their whole distributions. We focus on predictability in distribution by asking whether ethanol returns can be used to forecast different parts of field crops returns distribution, or vice versa. Density forecasts are constructed using Conditional Autoregressive Expectile models estimated with Asymmetric Least Squares. Forecast evaluation relies on quantile-weighted 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 distribution.

---

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

---

<sup>\*</sup> University of Milan-Bicocca and Fondazione Eni Enrico Mattei, Milan, Italy.

<sup>†</sup> University of Milan and IEFE-Bocconi, Milan, Italy.

<sup>‡</sup> Corresponding Author: Matteo Manera, Department of Economics, Management and Statistics (DEMS), University of Milan-Bicocca, Via Bicocca degli Arcimboldi, 8, I-20126, Milan, Italy. E-mail: matteo.manera@unimib.it.

*Acknowledgements:* Previous versions of this paper have been presented at: 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; the Department of Economics, Quantitative Methods and Management (DEMM), University of Milan; the Department of Economics, Management and Statistics (DEMS), University of Milan-Bicocca. The first author gratefully acknowledges the following research grant: “Dote Ricercatori: FSE, Regione Lombardia”.

# Food versus Fuel: Causality and Predictability in Distribution

## 1. Introduction

The dynamics of field crops prices has been the subject of broad attention from the media, public opinion, as well as the scientific community in recent years. 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. In addition to the natural causes of food price rises, such as bad harvests due to drought and floods, 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.

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. In particular, the Environmental Protection Agency's 2005 Renewable Fuel Standard and the 2007 Energy Independence and Security Act (EISA) are worth mentioning, according to which fuel blenders are obliged to mix a given amount of eligible biofuels into gasoline. In 2011, the mandate for corn-based biofuels under EISA is 12.6 billion gallons, which increases to 15 billion gallons in 2015. These figures imply that 39% of U.S. field corn is used to produce ethanol. Last year, due to a severe drought affecting more than 65% of farmland in the continental U.S., nearly half the corn crop has been directed to produce ethanol.<sup>1</sup>

In this paper we examine the causal nexus between ethanol and corn and other agricultural commodities, seeking support in favor or against the "Food versus Fuel" claim that food price inflation is primarily due to the ethanol production boom in the U.S..

Our analysis studies the relationship between the price of ethanol and the price of field crops and cattle in Nebraska from January 1987 through March 2012.<sup>2</sup> We concentrate on Nebraska for three main reasons. First, a practical motivation: a lengthy monthly time series of ethanol prices is freely available from the Nebraska Energy Office. Given that forecast analysis

---

<sup>1</sup> See, among others, Rosegrant et al. (2008), Mercer-Blackman et al. (2008), Mitchell (2008), UNCTAD (2008), Weise (2011), Parker (2013), The Economist (2013).

<sup>2</sup> As of February 2011, 27 states in the U.S. had operating ethanol facilities: Iowa, Nebraska and Illinois have a nameplate capacity equivalent to 26.10%, 12.97% and 9.02%, respectively, of the nation's total (13596 million gallons per year. Source: Nebraska Energy Office).

requires splitting the sample into an estimation and evaluation period, a sufficiently wide data coverage is necessary to obtain accurate estimates and to guarantee that out-of-sample tests have reasonable power. Second, the relevance of Nebraska production in the U.S. ethanol market (in 2011 it ranked second both in terms of nameplate capacity, 1764 million gallons per year, and operating production, 1739 million gallons per year) and within the state corn market (the Nebraska Corn Board has estimated that in the 2010-11 marketing year 35% of Nebraska's corn production was transformed in ethanol). Moreover, according to Solomon et al. (2007), four of the leading ethanol producing firms in the U.S. have distilleries in the state of Nebraska. Third, several studies have confirmed the importance of Nebraska for the purpose of studying the biofuel-food relationship. In particular, Serra et al. (2011) and Blomendahl et al. (2011) have used Nebraska data to study nonlinearities in the U.S. corn-ethanol-oil-gasoline price system and the impact of ethanol plants on surrounding farmland values. Wixson and Katchova (2012) have tested for asymmetric price transmissions in the grain and energy markets by relying on the price of ethanol in Nebraska. Elobeid and Tokgoz (2008) have calibrated their model of the U.S. ethanol market using Nebraska price data, as fully representative of the characteristics of the U.S. ethanol market.

Many studies have analysed the impact of biofuels on commodity food prices. As pointed out by Zilberman et al. (2012), there are two main strands in the literature on biofuels. The first relies on time-series econometrics to analyse the linkages between biofuel and food prices. The second, by means of simulation- and theory-based methods, deals with the impact of the introduction of biofuels on food prices. 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. Simulation-based analyses highlight that the introduction of biofuels may affect food prices and that this effect varies across regions and crops.

Our paper can be placed in the first strand of the literature.

We use time-series methods to analyse causal linkages between returns on ethanol and field crops in the U.S. by considering the whole distribution of returns, rather than focusing on few specific moments such as the mean or the variance. This new perspective of investigation is appropriate for at least two reasons. First, since returns are generally non-normal, their distribution can be hardly summarized by the mean. Second, rational decision makers rely on density forecasts in order to maximize their expected utility functions.

We focus on out-of-sample relations to answer the following questions: *a)* Can lagged returns on ethanol be used to forecast field crops 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. 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, findings *ii*), *iii*), *v*), *vi*) and *vii*) represent fresh new evidence on the biofuels-food relation, which is not supportive of the “Food versus Fuel” claim.

The rest of the paper is organized as follows. In the next Section we briefly review the literature. In Section 3 we motivate our modelling choices and provide economic justifications for the interpretation of results. The dataset is presented in the Section 4, while the methodology is illustrated in Section 5. Section 6 contains the empirical results and Section 7 concludes.

## 2. Related Literature

The “Food versus Fuel” claim 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 causality links. Given the approach followed in our paper, in this section we concentrate on contributions pertaining to the second strand of research.<sup>3</sup>

The studies testing the presence and the direction of the relationship between fuel and agricultural prices deal with a variety of empirical methods (e.g. structural vs. reduced form models, linear vs. non-linear models, statistical vs. econometric methods) applied to weekly or monthly spot and futures prices. 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.

---

<sup>3</sup>We address the reader interested in a broader survey of the literature to Zilberman et al. (2012).

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.

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.

Kristoufek et al. (2012b) rely on weekly price data for the period between November 2003 and February 2011 to analyze relations between biofuels (ethanol and biodiesel), their production factors (corn, wheat, soybeans and sugarcane) and fossil fuels (Brent crude oil, German diesel and U.S. gasoline). Short-run Granger causality is found running from corn prices to ethanol prices.

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 analyze 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) analyzes 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) analyze 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.

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 on the “Food versus Fuel” debate 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. Modelling Approach**

The line of investigation of the “Food versus Fuel” claim followed in this paper relies on the implicit assumptions that causality runs from ethanol price to corn prices and from corn prices to the price of other corn-based products, and that the amount of arable land is fixed over the short run.<sup>4</sup>

---

<sup>4</sup> See Abbott (2012) and Anderson et al. (2008). Different assumptions characterize the contributions by Zilberman et al. (2012), who provide a conceptual justification of the reason why causality may run from food to biofuel prices within a partial equilibrium framework, and by Piroli et al. (2012) who, from a long-run perspective, study the impacts of changes in biofuel prices on land use changes.

Moreover, our empirical approach is very different from previous works focusing on the linkages between ethanol and agricultural commodity markets, at least in two respects.

First, we use out-of-sample Granger causality tests. A common feature to most of the empirical literature is to analyze the relationship between the price of ethanol and the price of corn or, more generally, biofuel prices and agricultural prices, using in-sample Granger causality tests. Nothing is said about the out-of-sample performance of the estimated models, although this approach is statistically more appropriate to investigate Granger causality among variables, as originally put forth by Granger (1969). Our paper fills this gap. We run different tests of predictive ability to compare the forecasting performance of alternative autoregressive models with exogenous variables (i.e. AR-X models), against some benchmark models with no covariates, such as the random walk with drift. The benchmark models impose the joint null hypothesis of no-predictability and no-Granger causality. Therefore, if an AR-X model leads to superior out-of-sample forecast performance with respect to the benchmark model, this evidence is interpreted as Granger causality running from the exogenous variable to the dependent variable in the AR-X regression.

Second, we model and forecast the whole distribution of returns on ethanol and field crops prices. The analysis of the whole distribution of returns, rather than focusing on few specific moments such as the mean or the variance, is informative for different reasons. First, since returns are generally non-normal, the mean is hardly a valid summary of their distribution. Second, a density forecast of the future outcomes of a variable represents an estimate of the probability distribution of the possible realizations of that variable over the forecasting horizon, hence it provides a natural measure of the uncertainty associated with its future predictions. Third, it can be shown that, with exception of few of special cases, rational decision makers rely on density forecasts in order to maximize their expected utility functions (see Granger and Pesaran, 2000). Point forecasts are justified only when agents face linear-quadratic (LQ) decision problems, whereas, for non-LQ problems, optimal decision rules depend on the whole predictive distribution (Pesaran and Skouras, 2002).<sup>5</sup>

We estimate density forecasts for returns on ethanol and field crops prices with Asymmetric Least Squares (ALS, Newey and Powell, 1987). ALS is similar to Ordinary Least Squares

---

<sup>5</sup> A classical example is the maximization of the expected utility of an investor wishing to allocate his/her wealth across risky assets. In this case, point forecasts of the mean and the variance are sufficient to solve the investor's problem only if his/her preferences can be characterized with a quadratic utility function and the returns' distribution is arbitrary or, when preferences are arbitrary, if returns are multivariate normally distributed (Huang and Litzenberger, 1998, p. 61). However, for more general preferences and returns' distributions, optimal asset allocation requires density forecasts of the returns (see Cenesizoglu and Timmermann, 2008).



(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”.<sup>6</sup>

Both expectiles and quantiles can be used to describe the distribution of a random variable. From a computational point of view, direct estimation of expectiles has a number of advantages over the quantile regression approach. In particular, quantile regression is based on an absolute error loss function (known as check, or lin-lin loss), which is not continuously differentiable, hence it requires linear programming methods (Kroenker and Hallock, 2001). Conversely, ALS estimates can be computed with a straightforward application of the Iterated Weighted Least Squares (IWLS) algorithm (Newey and Powell, 1987; and Efron, 1991). Moreover, although the quantile estimator is generally more robust to outliers than the ALS estimator, Newey and Powell (1987) have shown that the quantile regression approach can be relatively inefficient for error distributions which are close to Gaussian.

Expectiles are less immediate to interpret than quantiles. For this reason, 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 estimated expectile regression lines. We then use the estimated quantiles in density forecasting to analyse the predictability of the distributions of returns on ethanol and field crops prices.

Expectiles and quantiles allow to investigate the predictability of returns in different regions of their distribution, such as the centre, one tail or both tails. This is an additional, innovative feature of our paper. As highlighted in Section 2, the majority of the contributions find evidence of Granger causality running from the prices of field crops to ethanol, but not vice versa. However, 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 their distributions. In our paper we extend the analysis to the whole distributions of returns, as well as to specific areas of the distributions other than the first and second moments, by estimating a sufficiently large number of quantiles with the ALS expectile approach. Our approach encompasses more traditional contributions, since the information provided by

---

<sup>6</sup> The expectile-based approach has already been successfully applied to volatility and density forecasts evaluation by Granger and Sin (2000), as well as in risk management analyses by Kuan et al. (2009) and Taylor (2008). Similarly, Cenesizoglu and Timmermann (2008) and Pedersen (2010) have studied the predictability of the distribution of stock returns using quantile regression. Isengildina-Massa et al. (2010, 2011) use quantile regression to construct confidence intervals for wheat, corn and soybean price forecasts issued by the U.S. Department of Agriculture. Lee and Yang (2012) study the Granger causality beyond the conditional mean between money and income by forecasting conditional quantiles. Galvao JR. et al. (2013) study quantile regression in an autoregressive dynamic framework with exogenous stationary covariates applied to house price returns in the U.K.. Candelon et al. (2013) use Granger causality in distribution tails to investigate oil markets integration.

different quantiles can also be used to retrieve specific sample moments. Specifically, location and volatility estimates can be derived from the median and the interquartile range (Pearson and Tukey, 1965), while skewness and kurtosis can be calculated using the methods suggested by Kim and White (2004).

In our paper the estimated density forecasts are evaluated with the quantile scoring rule proposed by Gneiting and Ranjan (2011). A scoring rule is a loss function for density forecasts, which associates a lower score to a better forecast. The authors have extended the quantile scoring rule by assigning more weight to the part of distribution (either centre, tails, right or left tail) which is of interest for the researcher. This “flexible” scoring rule is relevant for a variety of forecasts users. For instance, the tails of the distribution are usually the main focus of risk managers, while policy makers, in order to obtain confidence intervals around point forecasts, are generally more interested in the centre of the distribution.

Finally, we implement the conditional predictive ability test of Giacomini and White (2006) to check the statistical significance of the difference between the scoring rule of the AR-X model and the scoring rule of the benchmark model (i.e. score function differential). We analyse both the unconditional and conditional performance of each model. Unconditional tests indicate which forecast is more accurate on average in the past, while conditional tests use available information to predict which forecast is more accurate in the future.

#### **4. Data**

Our dataset comprises five monthly time series of nominal spot prices, namely ethanol, corn, soybeans, wheat and cattle, recorded in Nebraska from January 1987 through March 2012 (December 2010 for cattle). 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 cattle price 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. Time-varying weights based on the dollar value of production of field crops and cattle, sourced from the USDA database, have been used to construct two price indices which summarize field crops and cattle price dynamics. The first index includes the three field crops prices, while the second index adds the price of cattle to prices included in the first price index.<sup>7</sup>

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

---

<sup>7</sup> More details about the dataset and the construction of indices are provided in Appendix A1.

**[Figure 1 about here]**

**[Table 1 about here]**

As shown in Figure 1, ethanol prices have 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, 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 change in prices (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.

## 5. Density Forecasts with Expectiles

### 5.1 Expectile models

Our forecasting strategy can be illustrated as follows. The variables of interest are ethanol (ETH), corn (COR), soybean (SOY), wheat (WHE), cattle (CAT), and two price indexes (PI1, PI2). Variables are indexed by the subscript  $i = \text{ETH, COR, SOY, WHE, CAT, PI1, PI2}$ . Percentage price variations (i.e. returns) on each variable are computed as  $r_{it} = 100 \times \ln(P_{it} / P_{it-1})$ , 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:

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

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

where  $j = \text{COR, SOY, WHE, CAT, PI1, PI2}$  and  $t = 2, \dots, T_0$ .

In models (1a)-(1b),  $\tau_j(\omega | \Omega_{t-1})$  and  $\tau_{ETH_t}(\omega | \Omega_{t-1})$ , for  $\omega \in (0, 1)$ , denote the 100 $\alpha$ th conditional expectile of returns on variable  $j$  and ethanol, respectively, whereas  $\Omega_{t-1}$  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.

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

Each model (1a) and (1b) (i.e. 12 models total) is estimated with ALS and quantiles are retrieved as the proportion of returns lying below the 100 $\alpha$ th fitted expectile curve. The ALS estimator is the solution of minimizing an asymmetric quadratic loss function of the residuals from models (1a)-(1b). Since this loss function is continuously differentiable in the residuals, the parameters of models (1a)-(1b) can be estimated with IWLS.<sup>8</sup>

Newey and Powell (1987) have shown that if data are independently and identically distributed, the ALS estimator is consistent and asymptotically normal. An extension of these results to stationary and weakly dependent data, which is relevant for our paper, has been provided by Kuan et al. (2009).

Since each model (1a)-(1b) is estimated to match expectiles with quantiles  $\alpha = 0.05, 0.10, 0.15, \dots, 0.90, 0.95$  (i.e. 19 quantiles), 228 series of one-period ahead forecasts are computed.<sup>9</sup>

The size of the estimation sample is  $T_0 = 180$ , which corresponds to 50% of the total number of observations,  $T$ , for the returns on ethanol. For each model and quantile, the vector of  $H$  forecasts is obtained with a rolling window procedure. We start by estimating each model using observations from  $t = 2$  to  $t = T_0$ , and calculate forecasts in  $t = T_0 + 1$ . Then, observations from  $t = 3$  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. The rolling window forecasting scheme satisfies the assumptions underlying the Conditional Predictive Ability (CPA) test of Giacomini and White (2006), whose asymptotic distribution is obtained for fixed  $T_0$  and  $H \rightarrow \infty$ . 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.

---

<sup>8</sup> We use OLS estimates as the starting values for the IWLS algorithm and set the convergence criterion equal to  $10^{-12}$ .

<sup>9</sup> This figure excludes the forecasts obtained using the benchmark models.

One problem that might affect expectile and quantile regression estimates and forecasts is the violation of the monotonicity condition. In other words, estimated conditional quantile and expectile curves might exhibit some “crossings” (Kroenker and Xiao, 2006). Although in our case the incidence of this potential drawback is negligible (i.e. there are only 58 crossings out of 39900 forecasts, which corresponds to less than 0.6% of the forecasts obtained with CARE-X models), we tackle the problem with an “insanity filter”, which corrects monotonicity violations by substituting crossing CARE-X forecasts with the following benchmarks (see also Swanson and White, 1995):

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

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

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  $100\alpha$ th 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). For instance, an in-sample test of Granger causality running from variable  $j$  to ethanol would involve testing the null hypothesis  $H_0: \gamma_1(\omega) = 0$  in model (1b). Conversely, testing the null hypothesis  $H_0: \beta_1(\omega) = 0$  in model (1a) corresponds to an in-sample test of Granger causality from ethanol to variable  $j$ . 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. If the lowest loss is associated to models (1a)-(1b), then we can conclude that returns on ethanol [variable  $j$ ] can be predicted using returns on variable  $j$  [ethanol].

## 5.2 Evaluating quantile and density forecasts

We evaluate forecasts by means of the asymmetric quadratic loss function:

$$L_{t,\alpha}(r_{it}, \hat{q}_{\alpha,t}) = [\alpha + (1-2\alpha) \times I(r_{it} - \hat{q}_{\alpha,t} \leq 0)] \times |r_{it} - \hat{q}_{\alpha,t}|^2 \quad (3)$$

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

For each quantile  $\alpha$ , a given CARE-X model produces more accurate forecasts than its benchmark if the average loss (defined as:  $L_{\alpha}(\cdot) = H^{-1} \sum_{t=T_0+1}^T L_{t,\alpha}(\cdot)$ ) of the CARE-X model is smaller than the average loss of the corresponding CE model.

The quantile scoring rule (QS) of Gneiting et al. (2011) is the second tool we use for forecast evaluation. While the asymmetric quadratic loss delivers a picture of the predictive performance of a model in different parts of the distribution (i.e. in different quantiles, taken in isolation), 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 can be written as:

$$QS_{\alpha}(r_{it}, \hat{q}_{\alpha,t}) = 2 \times [I(r_{it} - \hat{q}_{\alpha,t} \leq 0) - \alpha] \times (\hat{q}_{\alpha,t} - r_{it}) \quad (4)$$

The QS (4) 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 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 proposed a weighted version of the QS given by:

$$WQS_{\alpha}(r_{it}, \hat{q}_{\alpha,t}) = v(\alpha) \times QS_{\alpha}(r_{it}, \hat{q}_{\alpha,t})$$

where  $v(\alpha)$  is a weight function that assigns a higher score to the desired part of the distribution. We consider four weighted quantile scores to focus on the following parts of the distribution: 1) Centre:  $v(\alpha) = \alpha(1 - \alpha)$ ; 2) Tails:  $v(\alpha) = (2\alpha - 1)^2$ ; 3) Left Tail:  $v(\alpha) = (1 - \alpha)^2$ ; 4) Right Tail:  $v(\alpha) = \alpha^2$ . Therefore we have used a total of five scoring rules for each CARE-X model and its corresponding benchmark. Noting that the uniform quantile score (4) is obtained for  $v(\alpha_k) = 1$ ,  $k = 1, \dots, K = 19$ , the average score for the  $i$ -th CARE-X model is:

$$AWQS_t^{CARE-X,i} = \sum_{k=1}^K v(\alpha_k) \times QS_{\alpha_k}(r_{it}, \hat{q}_{\alpha_k,t}) \quad (5)$$

for  $t = T_0 + 1, \dots, T$ . The  $i$ -th *CARE-X* model is on average more accurate at forecasting the whole distribution of returns on variable  $i$ , or some parts of it (e.g. centre, one or both tails), than its corresponding *CE* (benchmark) model if:  $\overline{AWQS}^{CARE-X,i} < \overline{AWQS}^{CE,i}$ , where  $\overline{AWQS}^{CARE-X,i} = H^{-1} \sum_{t=T_0+1}^T AWQS_t^{CARE-X,i}$ .

The CPA statistic of Giacomini and White (2006) tests the significance of score differentials,  $\Delta S_t$ , where  $\Delta S_t = AWQS_t^{CARE-X,i} - AWQS_t^{CE,i}$ . The CPA test is very flexible, since it applies to point, interval, probability or density forecasts for a wide range loss functions and data-generating processes, while forecasts can be obtained from nested and nonnested models estimated with parametric, semi-parametric, nonparametric, as well as Bayesian methods.

The CPA test is a Wald-type test statistic that can be written as:

$$W = H \left( H^{-1} \sum_{t=w+1}^T h_t \Delta S_t \right)' \Lambda^{-1} \left( H^{-1} \sum_{t=w+1}^T h_t \Delta S_t \right) \quad (6)$$

where  $h_t$  is a  $q \times 1$  vector of test functions,  $\Lambda$  is an heteroscedasticity- and autocorrelation-consistent estimate of the asymptotic covariance matrix of  $h_t \Delta S_t$ , and the null hypothesis to be tested is  $H_0: E(h_t \Delta S_t) = 0$ , for  $t = T_0 + 1, \dots, T$ . Under the assumption that forecasts originate from “limited memory methods”, such as rolling windows or the fixed estimation sample forecasting scheme (according to which the size of the estimation sample,  $T_0$ , is kept constant, while the size of the evaluation sample,  $H$ , increases), the CPA test (6) tends to a  $\chi^2(q)$ , as  $H \rightarrow \infty$ .

For  $q = 1$  and  $h_t = 1$ , the CPA test collapses to an unconditional test of predictive ability. In this 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. 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.<sup>10</sup> For this reason, we set  $q = 2$  and specify  $h_t$  as a function of a constant term and the NOPI dummy variable. A rejection of the null hypothesis provides evidence that the forecast 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.

## 6. Empirical Results

### 6.1 In-sample results

Estimates of CARE-X models (1a)-(2a) for the period January 1987 through March 2012 (December 2011 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

---

<sup>10</sup> Net oil price increases proxy oil shocks. Following Hamilton (1996), we construct a dummy variable, NOPI, 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. The test function  $h_t$  includes  $NOPI_{t-1}$ .



commodity showing some statistically significant coefficients associated to lagged returns on ethanol.

In summary, we are unable to find any empirical evidence of a solid 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 corn and field crops.

The pattern of coefficients reported in Panel b) of Table 2 is completely different. In CARE-X models with returns on ethanol as the 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.

## **6.2 Forecast Evaluation**

The evaluation period for one-month ahead forecasts spans April 1997-March 2012 ( $H = 180$ ) for PI1 and field crops and April 1997-December 2010 ( $H = 165$ ) for PI2 and cattle.

For each quantile and model, we compute the asymmetric quadratic loss (3), which 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 3 we show the asymmetric quadratic loss for CARE-X models (1a), where (lagged) ethanol is the exogenous explanatory variable and returns on variable  $j$  is the dependent variable. It is important to remember that ethanol has no in-sample predictive power for both price indices, cattle and most field crops. 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. 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 than 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.

**[Table 3 about here]**

Panel (b) of Table 3 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. Only in correspondence to the 0.90 and 0.95 quantiles do the CE models perform best, 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 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,  $i = \text{ETH, COR, SOY, WHE, CAT, PI1, PI2}$ . As shown by Elliott and Timmermann (2004), when the loss function is specified as equation (3), the optimal forecast combination weights can be estimated with IWLS. If the optimal combining weight of the forecasts obtained with CARE-X model (1a) ( $\phi_1$ ) is equal to one and the optimal combining weight of the forecasts obtained with CE model (2a) ( $\phi_2$ ) 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). Analogously, if the optimal combining weight of the forecasts obtained with CARE-X model (1b) ( $\psi_1$ ) is equal to one and the optimal combining weight of the forecasts obtained with CE model (2b) ( $\psi_2$ ) 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  $\phi_1=1$  ( $\psi_1=1$ ) and  $\phi_2=0$  ( $\psi_2=0$ ) are reported in Table 4. 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  $\phi_1=1$  and  $\phi_2=0$  are never rejected.

**[Table 4 about here]**

The results reported in Panel (b) have a penchant for “forecast encompassing”. In the case of PI1 and field crops, the null hypotheses  $\psi_1=1$  and  $\psi_2=0$  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.

So far, we have analysed the forecasting performance across different quantiles taken in isolation. Our investigation has highlighted that: (i) while ethanol cannot be used to forecast any part of the distribution of returns on price indices, corn, wheat and cattle, (ii) nevertheless it has some predictive power for soybean; (iii) field crops can be used to forecast the centre and the left tail of the distribution of ethanol returns.

We use the quantile score function and the CPA test to evaluate the predictive performance of each model on the whole distribution, that is jointly across quantiles.

The weighted versions of the quantile scoring rules (4)-(5) 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 5.

**[Table 5 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 density forecasts from the benchmark CE model (2b). However, the UPA test indicates that the score differential is not statistically different from zero. Two are the CPA tests reported in Table 5. The first, CPA1, includes in the test function of equation (6) 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 invariably 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 performance 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. Focussing our attention to 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).

## **7. Conclusions**

In this paper we have analyzed the predictive relationship between returns on ethanol, field crops and cattle in the state of Nebraska. The ultimate aim is to seek support in favor or against the “Food versus Fuel” debate, according to which food price inflation is primarily due to the ethanol production boom in the U.S.. An important implication of this view is that

causality should run from the price of ethanol to the price of corn and, by extension, other field crops.

Our analysis has been carried out on monthly data, spanning January 1987-March 2012. Density forecasts obtained by estimating different specifications of Conditional Autoregressive Expectile (CARE) models 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 percentage price variations (i.e. 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 and field crops returns can be obtained only by looking at their entire distributions.

In this paper, instead of focusing on specific moments of the distribution of returns, we have analysed the whole distribution of returns. This new perspective of investigation is appropriate for at least two reasons. First, since returns are generally non-normal, their distribution can be hardly summarized by the mean. Second, rational decision makers rely on density forecasts in order to maximize their expected utility functions.

We have calculated density and quantile forecasts using CARE models with exogenous variables 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?

Overall, our results do not support the general view of the “Food versus Fuel” debate.

More specifically, 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 Granger causes field crops or cattle. Rather, we provide empirical support for the existence of reverse causality, running from field crops to ethanol.

The novelty of our approach is twofold: *i)* we have carried out systematic forecast evaluation; *ii)* instead of focusing only on returns first and second moments, we have analysed the predictive relationships between ethanol and field crops for different quantiles taken in isolation and for their distributions as a whole.

Out-of-sample results are consistent with in-sample evidence. First, 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. Finally, the information content of returns on field crops can be fruitfully exploited to forecast extreme price decreases, which are of interest in risk management to compute value-at-risk or expected shortfall, as well as to construct prediction intervals for policy evaluation exercises.

## References

- Abbott, P. (2013), "Biofuels, Binding Constraints and Agricultural Price Volatility", forthcoming in J.P. Chavas, D. Hummels, and B. Wright (eds.), *The Economics of Food Price Volatility*, Chicago: University of Chicago Press
- Anderson, D., Anderson, J.D. and Sawyer, J. (2008), "Impact of the Ethanol Boom on Livestock and Dairy Industries: What Are They Going to Eat?", *Journal of Agricultural and Applied Economics* 40, 573-579
- Blomendahl, B.H., Perrin, R.K. and Johnson, B.B. (2011), "The Impact of Ethanol Plants on Surrounding Farmland Values: a Case Study", *Land Economics* 87, 223-232
- Candelon, B., Joëts, M. and Tokpavi, S. (2013), "Testing for Granger Causality in Distribution Tails: An Application to Oil markets Integration", *Economic Modelling* 31, 276-285
- Cenesizoglu, T. and Timmermann, A. (2008), "Is the Distribution of Stock Returns Predictable?", Working Paper, University of California at San Diego
- Efron, B. (1991), "Regression Percentiles Using Asymmetric Least Squared Error Loss", *Statistica Sinica* 1, 93-125
- Elliott, G. and Timmermann, A. (2004), "Optimal Forecast Combinations under General Loss Functions and Forecast Error Distributions", *Journal of Econometrics* 122, 47-79
- Elobeid, A. and Tokgoz, S. (2008), "Removing Distortions in the U.S. Ethanol Market: What Does It Imply for the United States and Brazil?", *American Journal of Agricultural Economics* 90, 918-932
- Galvao JR., A.F., Montes-Rojas, G. and Park, S.Y. (2013), "Quantile Autoregressive Distributed Lag Model with an Application to House Price Returns", *Oxford Bulletin of Economics and Statistics* 75, 307-321
- 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
- Huang, C.-F. and Litzenberger, R.H. (1998), *Foundations for Financial Economics*, Prentice Hall: New Jersey



- Isengildina-Massa, O., Irwin, S.H. and Good, D.L. (2010), “Quantile Regression Estimates of Confidence Intervals for WASDE Price Forecasts”, *Journal of Agricultural and Resource Economics* 35, 545-567
- Isengildina-Massa, O., Irwin, S.H., Good, D.L. and Massa, L. (2011), “Empirical Confidence Intervals for USDA Commodity Price Forecasts”, *Applied Economics* 43, 3789-3803
- Kim, T.-H. and White, H. (2004), “On More Robust Estimation of Skewness and Kurtosis”, *Finance Research Letters* 1, 56-73
- 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
- Koenker, R. and Hallock, K.F. (2001), “Quantile Regression”, *Journal of Economic Perspectives* 15, 143-156
- Koenker, R. and Xiao, Z. (2006), “Quantile Autoregression”, *Journal of the American Statistical Association* 101, 980-990
- 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”, CAMA Working Paper n. 38/2012
- Lee, T.-H. and Yang, W. (2012), “Money-Income Granger-Causality in Quantiles”, in D. Terrell and D. Millimet (eds.), *30<sup>th</sup> Anniversary Edition (Advances in Econometrics, Volume 30)*, Emerald Group Publishing Ltd, 385-409
- 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
- Newey, W.K. and Powell, J.L. (1987), “Asymmetric Least Squares Estimation and Testing”, *Econometrica* 55, 819-847
- 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>
- Pearson, E.S. and Tukey, J.W. (1965), “Approximate Means and Standard Deviations Based on Distances between Percentage Points of Frequency Curves”, *Biometrika* 52, 533-546
- Pedersen, T.Q. (2010), “Predictable Return Distributions”, *CREATES Research Papers 2010-38*, School of Economics and Management, University of Aarhus
- 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
- Piroli, G., Ciaian, P. and Kancs, d’A. (2012), “Land Use Change Impacts of Biofuels: Near-VAR Evidence from the US”, *Ecological Economics* 84, 98-109

- Rosegrant, M.W., Zhu, T., Msangi, S. and Sulser, T. (2008), "Global Scenarios for Biofuels: Impacts and Implications," *Review of Agricultural Economics*, 30, 495-505
- 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., 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
- Solomon, B.D., Barnes J.R. and Halvorsen, K.E. (2007), "Grain and Cellulosic Ethanol: History, Economics, and Energy Policy", *Biomass and Bioenergy* 31, 416-425
- Swanson, N. and White, H. (1995), "A Model-Selection Approach to Assessing the Information in the Term Structure Using Linear Models and Artificial Neural Networks", *Journal of Business and Economic Statistics* 13, 265-275
- Taylor, J.W. (2008), "Estimating Value at Risk and Expected Shortfall Using Expectiles", *Journal of Financial Econometrics* 6, 231-252
- The Economist (2013), "Fields of Gold", February 23<sup>rd</sup>, 37
- 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](http://www.unctad.org/en/docs/osg20081_en.pdf)
- Weise, E. (2011), "Ethanol Pumping Up Food Prices", USA TODAY.com, February, 14. Available at: [http://usatoday30.usatoday.com/money/industries/food/2011-02-09-corn-low\\_N.htm](http://usatoday30.usatoday.com/money/industries/food/2011-02-09-corn-low_N.htm)
- Wixson, S. E. and Katchova, A.L. (2012), "Price Asymmetric Relationships in Commodity and Energy Markets", mimeo. Paper presented at the 123<sup>rd</sup> 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. (2012), "The Impact of Biofuels on Commodity Food Prices: Assessment of Findings", *American Journal of Agricultural Economics*, forthcoming

## Tables & Figures

**Table 1.** Descriptive Statistics: January 1987 - December 2011/March 2012

Panel (a): 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. The time period spanned by the monthly nominal spot price of CAT and PI2 is January 1987-December 2011, 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 2011/March 2012**

Panel a) Does ETH Granger Cause variable $j$ ?						
$\alpha$	$j$					
	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 $j$ Granger Cause ETH?						
$\alpha$	$i$					
	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 in 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 of the table are the  $p$ -values for 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  $\alpha$  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. Asymmetric Quadratic Loss**

Panel a) CARE-X model (1a) vs CE model (2a)										
$j/\alpha$	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)										
$j/\alpha$	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  $\alpha$  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  $\alpha$ . In Panel b) an asterisk identifies the best model for each quantile  $\alpha$ .

**Table 4. Optimal Combining Weights: CARE-X and CE Forecasts**

Panel a)														
<i>j</i>	$\alpha$													
	0.05		0.10		0.25		0.50		0.75		0.90		0.95	
	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)														
<i>j</i>	$\alpha$													
	0.05		0.10		0.25		0.50		0.75		0.90		0.95	
	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 of 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  $\alpha$  indicate the quantiles estimated from expectiles. Numbers reported in Panel a) are the combining weights  $\phi_1$  and  $\phi_2$  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  $r_{jt}^{\text{CE}}$  are forecast from CE model (2a). Coefficients  $\phi_0$ ,  $\phi_1$  and  $\phi_2$  are estimated with Iterated Weighted Least Squares. If the single null hypotheses  $\phi_1=1$  and  $\phi_2=0$  are not rejected, then forecasting with CARE-X model (1a) is more accurate than forecasting with CE model (2a). \* (\*\*) [\*\*\*] denotes rejection of each single null hypothesis at 0.10 (0.05) [0.01] significance level. 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  $\psi_1$  and  $\psi_2$  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  $r_{ETHt}^{\text{CE}}$  are forecast from CE model (2b). Coefficients  $\psi_0$ ,  $\psi_1$  and  $\psi_2$  are estimated with Iterated Weighted Least Squares. If the single null hypotheses  $\psi_1=1$  and  $\psi_2=0$  are not rejected, then forecasting with CARE-X model (1b) is more accurate than forecasting with CE model (2b). \* (\*\*) [\*\*\*] denotes rejection of each single null hypothesis at 0.10 (0.05) [0.01] significance level.

**Table 5.** Density Forecasts: Score Function Differentials and CPA TestsPanel a) Does ETH help forecasting variable  $j$ ?

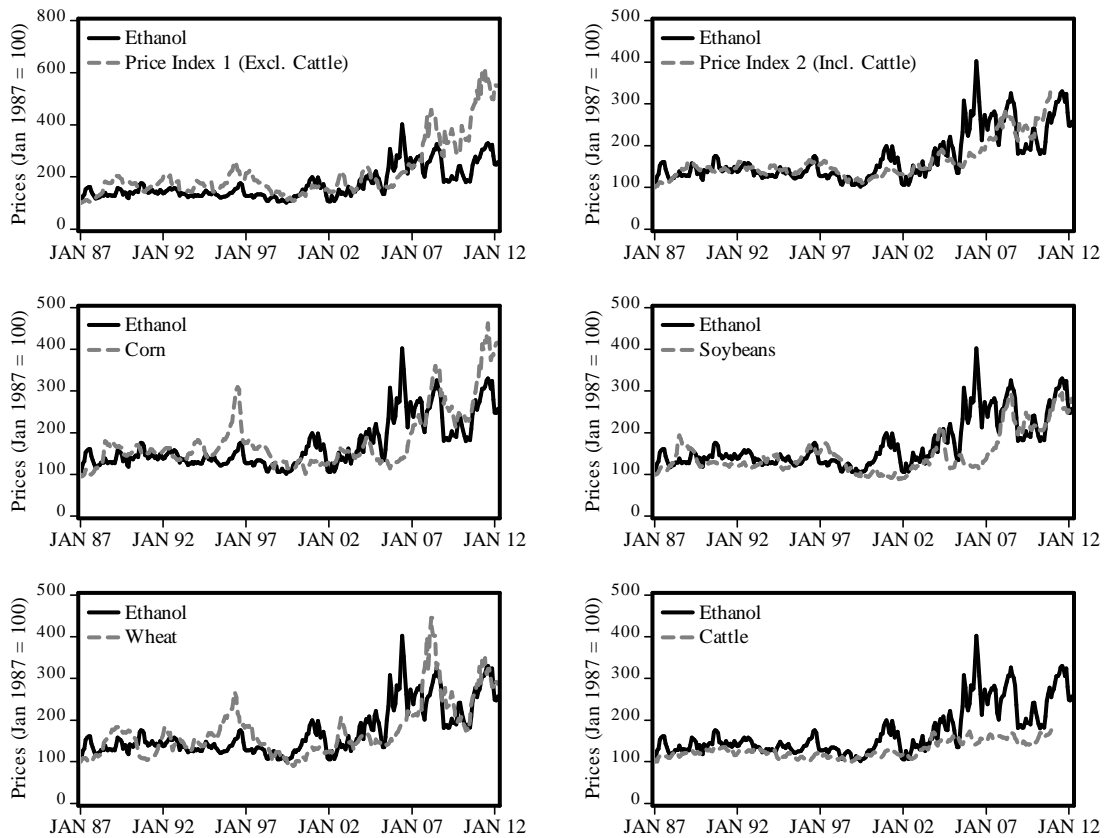
$j$	Weights: Uniform			Weights: Center			Weights: Tails			Weights: Left Tail			Weights: Right Tail		
	UPA	CPA1	CPA2	UPA	CPA1	CPA2	UPA	CPA1	CPA2	UPA	CPA1	CPA2	UPA	CPA1	CPA2
PI1	0.023	0.023	0.023	0.014	0.014	0.014	0.049	0.049	0.049	-0.032	-0.032	-0.032	0.080	0.080	0.080
PI2	-0.016	-0.016	-0.016	-0.025	-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 ETH?

$j$	Weights: Uniform			Weights: Center			Weights: Tails			Weights: Left Tail			Weights: Right Tail		
	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 of Table 1. This table reports standardized quantile weighted score function differentials ( $\Delta S$ ), as defined in Section 5.2. 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 unweighted distribution. In Panel a) negative numbers (i.e.  $\Delta S < 0$ ) indicate that the score function of CARE-X model (1a) is on average lower than the score function 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, when  $q=1$  and  $h_t=1$  for any  $t$ . CPA1 is the CPA test when  $q=2$  and  $h_t=(1, \Delta S_{t-1})$ . CPA2 is the CPA test when  $q=2$  and  $h_t=(1, \text{NOPL}_{t-1})$ , as defined in note 16.

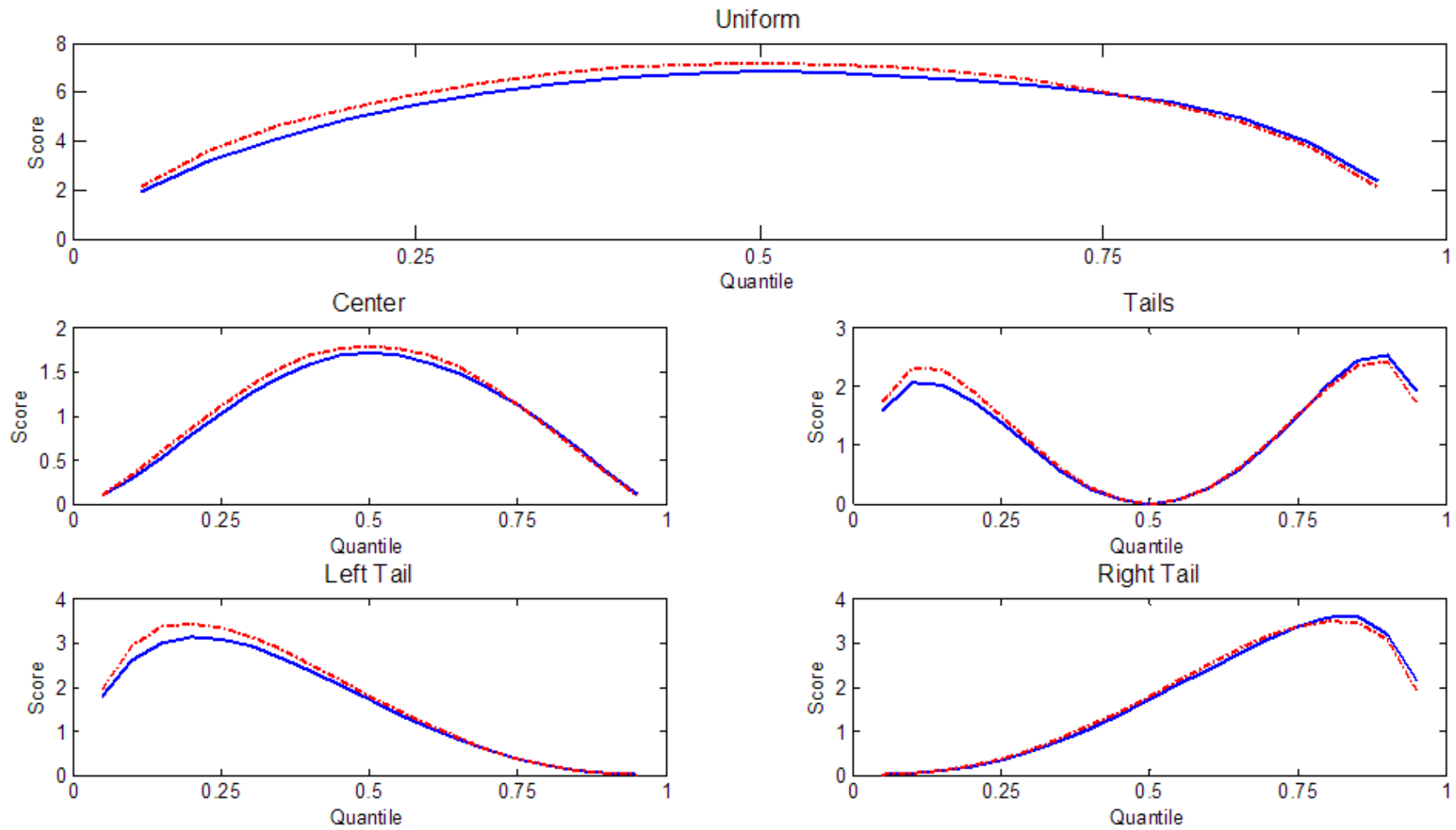
**Figure 1.** Prices: Ethanol, Indices, Field Crops and Cattle (current dollars)



*Notes:* All prices are represented on a common scale (i.e. 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

**Table A1.** Data description

Series ID	Description	Unit	Frequency	Time Period	Source <sup>a</sup>
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 - Dec/2010	USDA
PB	Cattle (>500 LBS): Price Received	Dollars per CWT <sup>b</sup>	Monthly	Jan/1982 - Mar/2012	USDA
YC <sup>c</sup>	Corn (Grain): Production	Dollars	Yearly	1982-2012	USDA
YS <sup>c</sup>	Soybeans: Production	Dollars	Yearly	1982-2012	USDA
YW <sup>c</sup>	Wheat: Production	Dollars	Yearly	1982-2012	USDA
YB <sup>c,d</sup>	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; BLS = Bureau of Labour Statistics; AC = Author's Calculations; (b) CWT = hundredweight; (c) The value for 2012 is obtained as a cubic trend forecast; (d) Yearly observations are obtained as averages of monthly data (i.e. CPIMW).

The production variables described in Table A1 have been used to construct two commodity price indexes. 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:

$$w_{j,1,t} = Y_{j,t} / (Y_{C,t} + Y_{W,t} + Y_{S,t}), \text{ for } j = C, S, W \text{ and } t = 1982, \dots, 2012 \quad (\text{A1})$$

$$w_{i,2,t} = Y_{i,t} / (Y_{C,t} + Y_{W,t} + Y_{S,t} + Y_{B,t}), \text{ for } i = C, S, W, B \text{ and } t = 1982, \dots, 2010 \quad (\text{A2})$$

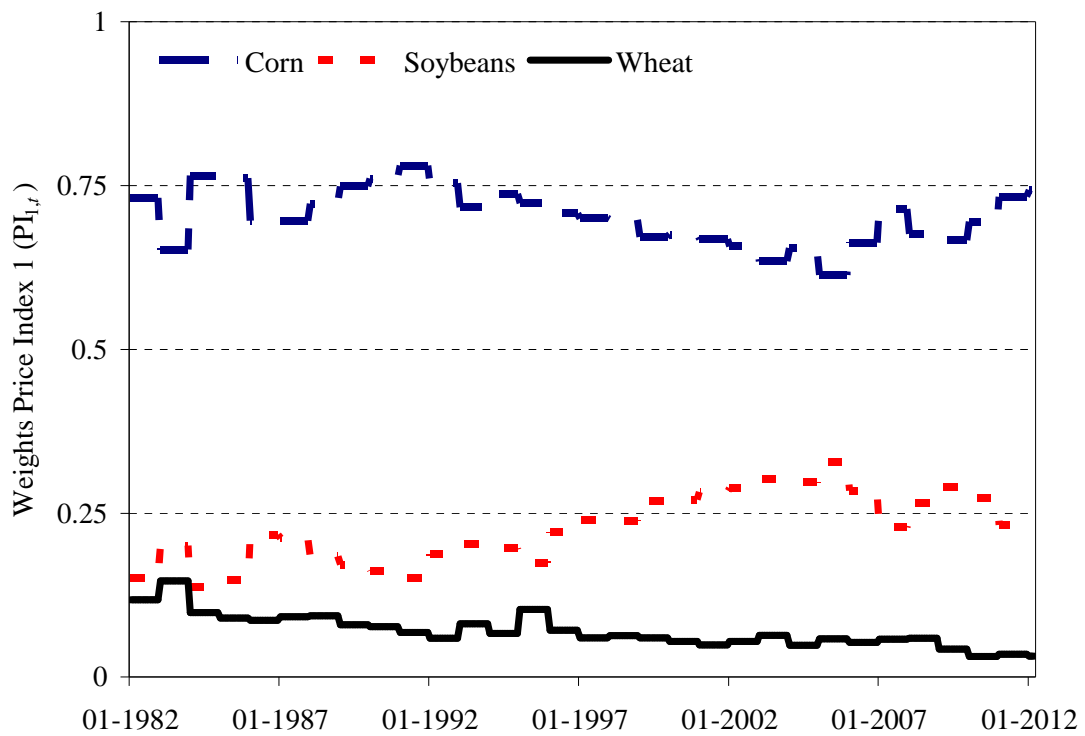
Given that production variables are recorded at yearly frequency, we constructed monthly observations by assuming constant weights within the year (e.g.  $w_{j,1,1/1982} = w_{j,1,2/1982} = \dots = w_{j,1,12/1982}$ , where  $r/1982$  indicates the  $r$ -th month of year 1982).

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

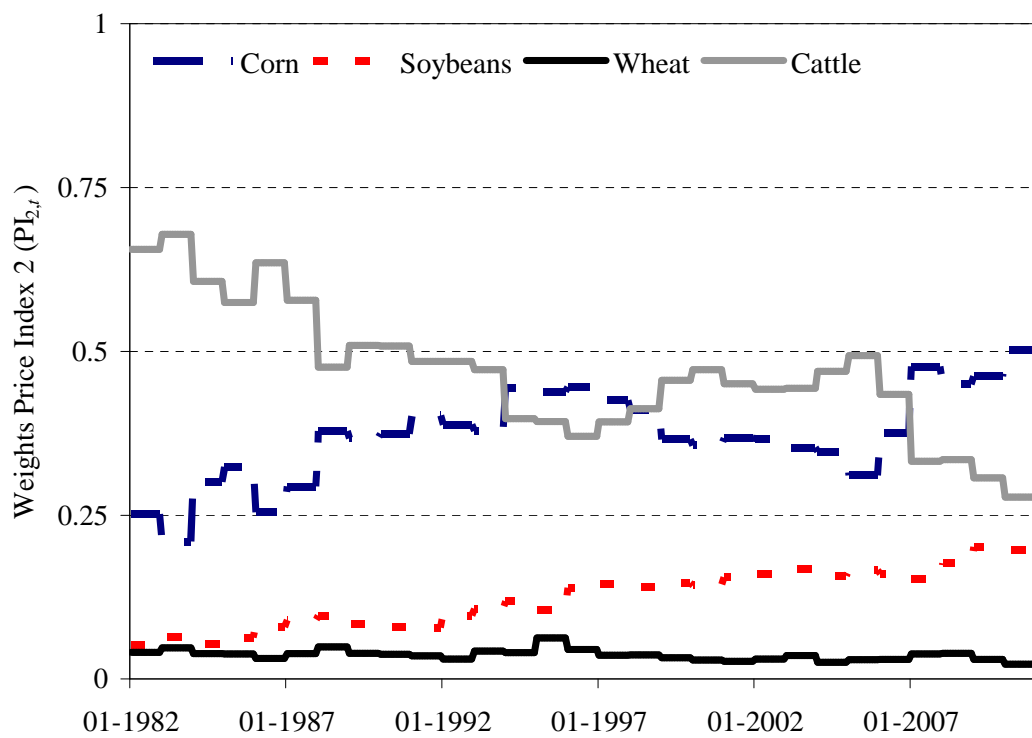
$$PI_{1,t} = (PC_t / w_{C,1,t}) + (PW_t / w_{W,1,t}) + (PS_t / w_{S,1,t}), \text{ for } t = 1/1982, \dots, 3/2012, \quad (\text{A3})$$

$$PI_{2,t} = (PC_t / w_{C,2,t}) + (PW_t / w_{W,2,t}) + \dots \\ + (PS_t / w_{S,2,t}) + (PB_t / w_{B,2,t}) \text{ for } t = 1/1982, \dots, 12/2010. \quad (\text{A4})$$

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



Panel a)



Panel b)