

The role of transition regime models for corn prices forecasting

O papel de modelos de regime de transição para previsão de preços do milho

Vinícius Phillipe de Albuquerque¹ , Rennan Kertly de Medeiros² , Diego Pitta de Jesus² , Felipe Araujo de Oliveira³ 

¹Programa de Pós-Graduação em Economia, Universidade Federal de Pernambuco (UFPE), Recife (PE), Brasil. E-mail: vinicius.phillipe@hotmail.com

²Programa de Pós-Graduação em Economia, Universidade Federal da Paraíba (UFPB), João Pessoa (PB), Brasil. E-mails: rennanmedeiros22@hotmail.com; diegopitta13@hotmail.com

³B3 S.A., São Paulo (SP), Brasil. E-mail: felipe.8.araujo@gmail.com

How to cite: Albuquerque, V. P., Medeiros, R. K., Jesus, D. P., & Oliveira, F. A. (2022). The role of regime transition models for corn prices forecasting. *Revista de Economia e Sociologia Rural*, 60(2), e236922. <https://doi.org/10.1590/1806-9479.2021.236922>

Abstract: Given the relevance of corn for food and fuel industries, analysts and scholars are constantly comparing the forecasting accuracy of econometric models. These exercises test not only for the use of new approaches and methods, but also for the addition of fundamental variables linked to the corn market. This paper compares the accuracy of different usual models in financial macro-econometric literature for the period between 1995 and 2017. The main contribution lies in the use of transition regime models, which accommodate structural breaks and perform better for corn price forecasting. The results point out that the best models are those which consider not only the corn market structure, or macroeconomic and financial fundamentals, but also the non-linear trend and transition regimes, such as threshold autoregressive models.

Keywords: forecasting, corn prices, accuracy, econometric models.

Resumo: Dada a relevância do milho para as indústrias de alimentos e combustíveis, analistas e acadêmicos estão constantemente comparando a precisão das previsões dos modelos econométricos. Esses exercícios testam não apenas o uso de novas abordagens e métodos, mas também a adição de variáveis fundamentais ligadas ao mercado de milho. Este artigo compara a precisão de diferentes modelos usuais na literatura macro-econômica financeira para o período entre 1995 e 2017. A principal contribuição está no uso de modelos de transição de regime, que acomodam quebras estruturais e têm melhor desempenho na previsão do preço do milho. Os resultados apontam que os melhores modelos são aqueles que consideram não apenas a estrutura do mercado de milho, ou fundamentos macroeconômicos e financeiros, mas também a tendência não linear e as transições de regime, como os modelos autorregressivos *threshold*.

Palavras-chave: previsão, preços do milho, acurácia, modelos econométricos.

1. Introduction

There is a growing need for a broader understanding of commodities price determinants and dynamics. Primary agricultural products such as corn play a dual role in this discussion. They are not just strongly affected by feed and food issues but also by strong growth of the global population, the need for sustainable practices of production of food, and for use in fuel complexes, mainly due to the rise in renewable energy use (Gurge, 2011).

According to Food and Agriculture Organization of the United Nations (2018) database, the world population is projected to reach almost 10 billion by 2050. Additionally, persistent poverty, unemployment, and inequality tend to restrain access to food and represent an obstruction to food security and nutrition goals, as well. Defeating poverty and inequality



impacts undernourishment positively, sparking demand for commodities, and consequently affecting their price dynamics.

Fuel commodities, such as corn, are sensible to shocks in the energy sector. As reported by the department of economic research of United States Department of Agriculture (2019), the corn crop channeled to ethanol production went from 32% to 45%, in 2018. This fact is more relevant if we consider that in 2005 only 10% of corn stocks were used for this purpose and in 2012, given corn prices upward trend, it reached 62%.

As this commodity is extremely demanded worldwide, it represents an important asset to primary-sector dependent countries. Thus, many authors explore the role of corn exports in the trade balance. For example, Alvim & Waquil (2005), who analyzed how trade agreements affected corn and other Brazilian grains in foreign trade. Also, Mariscal & Powell (2014), who investigated how booms and breaks impact trade balance long-run trends of developing countries. Also, Winkelried (2018) who analyzed the Presbisch-Singer hypothesis of a secular decline in commodity prices with time-series econometrics.

The precision of commodities prices is critical in the decision-making process of policymakers in their interventions in world trade, economic growth and, sectoral, and sustainability policies. Hitherto, the recent literature is concerned with variables or methods which promote the best accuracy in forecasting models.

Studies tried to improve forecasting performance by modeling price dynamics with different assumptions. For example, Ahumada & Cornejo (2016) explored the interdependence of corn, soy, and wheat using a Vector Error Correction Model (VECM) and, Zhang et al. (2018) used a quantile regression and a neural network to soy market in China. Winkelried (2016) decomposed food prices with an L1 filter. Favro et al. (2015) investigated correlation and causality between corn prices, soybean prices, and poultry farming.

Regarding the relationship between commodities and financial variables, Cuaresma et al. (2018) highlights is the importance of studying the impact of primary sector price index volatility and stock market volatility by focusing on the Arabic coffee case.

Outlining corn markets, the literature is interested in two approaches. The first one is attentive to volatility forecasting. Benavides (2009) assess it comparing implicit volatility with time series models. McPhail et al. (2012) and Serra & Gil (2013) attacks it in a more structural approach. They shed light on the role of global demand, speculation and energetic necessity, and macroeconomic conditions, as well. The second approach focuses on accessible but precise models to price forecasting to practitioners, as traders and regulators or producers. Bastianin et al. (2014) points out the economic relationship and time precedence of ethanol and corn. Hoffman et al. (2015), checks the quality of *World Agriculture Supply and Demand Estimates* - WASDE, estimates of the American agriculture department. Jadhav et al. (2017) used ARIMA models in corn price forecasting. Xu (2018) analyzes the co-integration bounded by corn prices of 182 spot markets. Osathanunkul et al. (2018) studies causality between oil and food commodities by a Bayesian Vector Auto-regressive Model (BVAR) and a Markov Switching BVAR (MS-BVAR).

In this sense, this paper aims to study the adherence and performance of several methods and models on corn price forecasting. It contributes to recent literature by accommodating non-linearity and structural breaks in time series using regime shift models. Our findings may help financial analysts, policymakers, and traders once it points out the use of econometric treatments to augment corn-price forecasting accuracy.

Beyond this introduction, the article is composed of four additional sections. The second presents the database and method. The third show data treatment. The fourth section declares the results. The last one presents a few conclusions.

2. Methodology and Data

To achieve our objectives, a mix of variables and models were used for corn price forecasting, then we compared their forecast accuracy measures by Root Mean Square Errors - RMSE, after using the Diebold-Mariano pairing method. The addition of micro and macroeconomic variables under different assumptions to improve the performance of corn-price forecasting required the evaluation of time-series dynamics.

The time series used are corn production, corn stocks, inflation, corn prices, and interest rates. We used monthly observations that goes from 1995 to 2017 (more details in Table 1). The data treatment follows Enders (2008), and it begins with the identification and removal of seasonality using X-13 Arima. Then, we tested for unit-root and structural breaks, with ADF and Zivot-Andrews tests, respectively. Additionally, we used the Box-Jenkins procedure for the correction of autocorrelation.

Table 1. Variables Summary

Variable	Source
Ethanol Production	US Department of Agriculture – USDA
Ethanol Stock	US Department of Agriculture – USDA
Ethanol Prices	Reuters
Inflation (Consumer Price Index)	Bureau of Labor Statistics – USBLS
Food Inflation	Food and Agriculture Organization of the United Nations – FAOUN
Cereals Price Index	Food and Agriculture Organization of the United Nations – FAOUN
Short-Term Interest Rates	FED
Long-Term Interest Rates	FED
Corn Spot Price	Index Mundi
Corn Production	US Department of Agriculture – USDA
Oil Production	Energy Information Administration – EIA Oil
Oil Spot Price	Investing Website

Source: Authors' elaboration.

According to Table 2, the following 13 econometric models were used: random walk, AR(p), MA(q), ARIMA(p,d,q), VARs, SETAR(p), STAR e LSTAR. Excluding Self-Exciting Threshold Autoregressive (SETAR), the other models were specified with a constant. Following Hotelling (1931) to accommodate effects on the level of these models. Specifications of estimated models are presented in Table 2.

VAR A model follows Reichsfeld & Roache (2011) and Xu (2018) acknowledging the price discovery function of futures contracts over spot prices. VAR B is specified considering corn market structure as highlighted in Gallagher (1986) and Jayne & Rashid (2010). VAR C aggregates VAR B variables casting energy demand, as in Trujillo-Barrera et al. (2012) and Mallory et al. (2012). VAR D sums to its specification some monetary variables, such as inflation, following Beckers & Beidas-Strom (2015) and Albuquerque et al. (2018). Models E and F specification investigates substitution effects on grains and cereals inflation, as proposed by McPhail et al. (2012) and Campos (2020).

Table 2. Estimated models specification

Model	Variables
Random Walk	Corn Spot Price
AR(p)	Corn Spot Price
MA(q)	Corn Spot Price
ARIMA(p,d,q)	Corn Spot Price
VAR A	Corn Spot Price and Future
VAR B	Corn Spot Price + Global Corn Production
VAR C	Model B + Ethanol Production + Ethanol Stock + Ethanol Production
VAR D	Model C + Inflation + Short and Long-term interest rates
VAR E	Model C + Food Inflation + Short and Long-term interest rates
VAR F	Model C + Grains Inflation + Short and Long-term interest rates
SETAR(p)	Corn Spot Price
LSTAR	Corn Spot Price
STAR	Corn Spot Price

Source: Authors' elaboration.

2.1 VAR, STAR and empirical strategy

Following Vector Autoregressive Models (VAR) introduced in Sims (1980), the problem of identification and autocorrelation of innovations in multivariate models can be surpassed when endogenous effects are considered. Algebraically it can be specified as in Equation 1.

$$Bx_t = \Gamma_0 + \Gamma_1 x_{t-1} + J_t \tag{1}$$

where B represents a matrix of coefficients, Γ_0 is a matrix of constants, Γ_1 the matrix of lagged variables ϵ_t are the white noise innovations of the model. Multiplying the Equation 1 by the inverse of matrix B (Equation 2):

$$B^{-1}Bx_t = B^{-1}\Gamma_0 + B^{-1}\Gamma_1 x_{t-1} + B^{-1} \epsilon_t \tag{2}$$

Simplifying 2 results in the VAR estimated Equation 3:

$$X_t = A_0 + A_1 X_{t-1} + \epsilon_t \tag{3}$$

where X_t represents the matrix of variables, A_0 a matrix of constants, A_1 is the matrix of estimated parameters, and ϵ_t represents the error term of this model. Therefore, the interdependence of the variables in VAR accommodates endogenous effects and autoregressive effects.

However, the model is an incomplete approach in the presence of non-linearities in the time series phenomenon to be studied. In that regard, STAR model extends VAR approach by introducing a state-space transition among regimes. The model was firstly proposed in Chan & Tong (1986).

Given the dependence of corn prices with global real economic activity, and the possibility of structural breaks, we tested transition regime models. As in Albuquerque et al. (2018), who suggested the SETAR model performed better than linear models for oil prices forecasting, we tested it.

For the identification and estimation of VAR models in this paper, we follow Enders (2008) and Lütkepohl (2005) steps: a) unit root tests to check for the order of integration regarding each variable; b) lags order identification relying on information criteria; c) cointegration identification using Johansen (1988); d) evaluation of autocorrelation, heteroskedasticity, and stability of the model; e) in-sample and out-of-sample forecasting; f) evaluation of forecasting performance using accuracy measures.

2.2 Forecasting Evaluation

Accuracy evaluation is schematized in two steps: i) comparing Root Mean Square Error to in-sample previsions; e b) using statistic proposed in Diebold & Mariano (2002) to out-of-sample forecasts.

Diebold-Mariano test infers over the difference of precision of two different models.¹ The null hypothesis is that the two forecasts have the same accuracy; whereas the alternative hypothesis is that two forecasts have different levels of precision.

The RMSE and Diebold-Mariano tests can be specified as in Equations 4 and 5:

$$RMSE_h = \sqrt{\frac{1}{T-h} \sum_{t=1}^{T-h} (\hat{y}_{t+h} - y_{t+h})^2} \quad (4)$$

$$DM_h = \frac{\bar{d}}{\sqrt{\frac{\hat{w}_d}{T-h}}} \quad (5)$$

where $\bar{d} = \frac{1}{T-h} \sum_{t=1}^{T-h} [(\hat{u}_{t+h}^B)^2 - (\hat{u}_{t+h}^A)^2]$ and $\hat{u}_{t+h}^A, \hat{u}_{t+h}^B$ are forecasting error to h-steps ahead in model A and B respectively, \hat{w}_d is the variance-covariance matrix \bar{d} , corrected to serial autocorrelation.

3. Time Series Analysis and Variables Treatment

This section presents a preliminary analysis of the variables used, and statistical treatment required to correct identification, as well. In this sense, subsection 3.1 explores the features of variables and period studied, and stylized facts. Subsection 3.2 presents the procedures and interpretation of variables treatment, using statistics to evaluate unit-roots, structural breaks, and cointegration.

3.1 Time Series Analysis

The time series of corn prices go from January/1995 to December/2017. In Figure 1 it can be observed the trends and series behavior. The period between 2000 and 2008 is marked by an accelerated uptrend in price, which is probably explained by a rising per capita income in India and China, hence stimulating higher domestic demand for this grain, and the ethanol market (production) in the U.S.

Notwithstanding, Baffes & Haniotis (2010) outlines that the rising consumption of this commodity by India and China is lower than global demand growth. Bobenrieth et al. (2004) and Runge & Senauer (2007) connect this price spiraling to a growing demand for corn ethanol by the USA. The latter is not a consensus in the literature, as Gilbert (2010) says that food price rises are more related to index-based investment in agricultural futures markets rather than to the above-mentioned factors.

¹ In its naive version considers symmetrical loss functions. In other words, it means that under and over-predict are equally weighted. In some cases, it could be inappropriate.

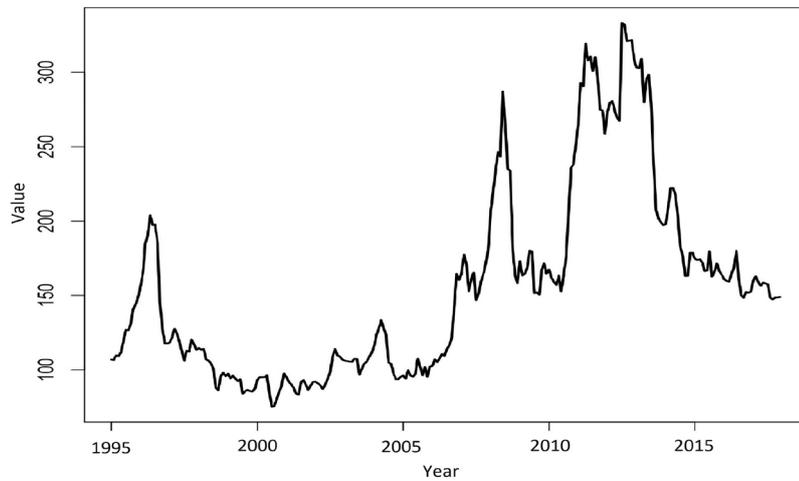


Figure 1. Time Series dynamics: Corn Price. Source: Authors' elaboration.

The decreasing shift in corn prices between 2008 and 2011 is largely explained by the subprime crisis. The aftermath is marked by a reaction of prices due to the low level of stocks. Roberts & Schlenker (2009) and Roberts & Schlenker (2013) shed light on the substitutability of corn on global market. The market structure of corn suggests that when stocks are extremely low, prices become highly sensitive to marginal shocks.

According to United States Department of Agriculture (2015), corn is the most produced grain in the world. The production is concentrated in three big players: the USA, China, and Brazil; representing 65% of global production. Thus, the drop in prices can be explained by rising stock levels in China. Figure 2 presents the trajectory of the main determinants of corn spot prices, as pointed out by the literature review.

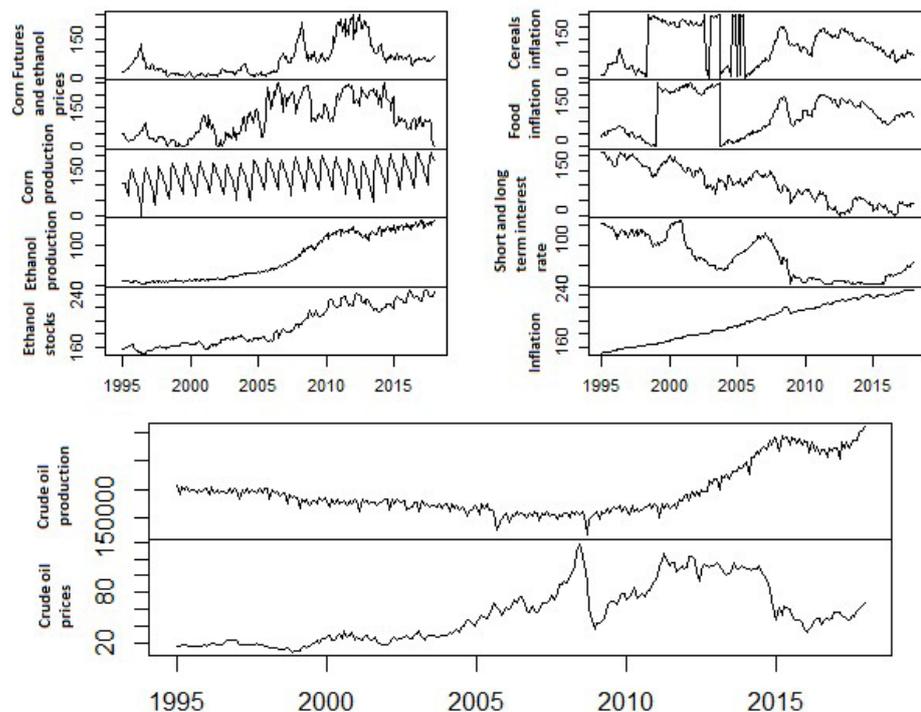


Figure 2. Determinants of corn prices. Source: Authors' elaboration.

According to Figure 2, all variables present a structural break due to the global financial crisis in 2008. Additionally, it can be observed a shift in the term structure of interest rates in the USA, corroborating with Krishnamurthy & Vissing-Jorgensen (2011) perception. Graphically, we can identify two main features to be treated: seasonality and trend.

3.2 Variables Treatment

Seasonality, if not correctly addressed can be dismal to forecasting. Hereof, it is used X-13 Arima, where the null hypothesis of the QS test is the non-existence of seasonal effects. The Table 3 presents these results.

Table 3. X13-ARIMA Seasonality Test

Variable	qsori	qsorievadj	qrsrd	qssadj	qssadjevadj
Ethanol Production	0.0000	0.0000	1.0000	1.0000	1.0000
Ethanol Stock	0.0000	0.0000	1.0000	1.0000	1.0000
Ethanol Spot Price	1.0000	1.0000	1.0000	1.0000	1.0000
Inflation (CPI)	0.0018	0.0001	1.0000	1.0000	1.0000
Food Inflation	1.0000	1.0000	1.0000	1.0000	1.0000
Grain Inflation	1.0000	1.0000	1.0000	1.0000	1.0000
Short-term Interest Rate	0.0310	0.0000	1.0000	1.0000	1.0000
Long-term Interest Rate	1.0000	1.0000	1.0000	1.0000	1.0000
Corn Spot Price	0.9205	0.4150	0.8632	0.9205	0.4150
Corn Future Price	1.0000	1.0000	0.9206	1.0000	1.0000
Corn Production	0.0000	0.0000	1.0000	1.0000	1.0000
Oil Production	0.0000	0.0000	0.9362	1.0000	1.0000
Oil Spot Price	0.8619	0.0314	0.9981	0.8619	0.0314

Source: Authors' elaboration.

Note: p-value is obtained by an approximation done by simulation of a qui-square probability density function with two degrees of freedom.

The null hypothesis of no seasonality is not accepted to ethanol production and stock, and inflation (CPI). These variables are important in this study due to their cross-dependence with the oil market.

Augmented Dickey-Fuller (ADF) test results are exposed to Table 4. Only short and long-term interest rates are stable in level, other ones present stochastic trend, being stationarity yielded by differencing the series.

Table 4. Unit Root Tests

Variable	Level	Difference	Lag
Ethanol Production	0.953	0.01	5
Stock of Ethanol	0.955	0.01	5
Spot Price of Ethanol	0.363	0.01	5
Inflation (CPI)	0.999	0.01	5
Food Inflation	0.774	0.01	5
Grains Inflation	0.636	0.01	5
Short-term Interest Rates	0.032	0.01	5
Long-term Interest Rates	0.033	0.01	5
Corn Spot Price	0.570	0.01	5
Corn Future Price	0.433	0.01	5
Corn Production	0.145	0.01	5
Oil Production	0.748	0.01	5
Oil Spot Price	0.536	0.01	5

Source: Authors' elaboration.

Note: p-value significance level of 1%.

Later on, is applied Zivot-Andrews to certify about structural breaks in time series. The results are presented in Table 5.

Table 5. Structural Break – Zivot Andrews

Variable	Potential Structural Breaks	t-calc	t-crit
Ethanol Production	set/2007	-5.3572	-5.08
Ethanol Stock	jul/2007	-4.1406	-5.08
Ethanol Spot Price	may/2005	-4.8198	-5.08
Inflation (CPI)	jan/2005	-3.7091	-5.08
Food Inflation	jun/2010	-2.6025	-5.08
Grains Inflation	jun/2010	-3.2547	-5.08
Short-term Interest Rates	oct/2007	-2.3335	-5.08
Long-term Interest Rates	apr/2011	-4.3918	-5.08
Corn Spot Price	jul/2010	-3.5652	-5.08
Corn Future Price	mar/2010	-4.5244	-5.08
Corn Production	nov/2011	-5.6320	-5.08
Oil Production	jul/2008	-7.8380	-5.08
Oil Spot Price	aug/2014	-3.7197	-5.08

Source: Authors' elaboration.
Note: Significance level of p-value is 5%.

Following Zivot & Andrews (2002), the break is mapped and endogenously estimated. The null hypothesis is of a unit root process with a potential break, the alternative hypothesis is the existence of a unique potential break in a trend stationary process. In Table 5, the Zivot-Andrews test infers to most of the variables a structural break in the aftermath of the subprime crisis in 2008. Just Inflation, ethanol spot price, and ethanol production seem to be decoupled from this event. This is expected as Inflation is the first-difference series. Also, the ethanol market suffered numerous influences with domestic demand and China, and India's demand.

SETAR model is used to accommodate structural breaks and non-linearities in the variables used. It follows the tradition of regime shift models, where a threshold separates levels of parameters. When the threshold is identified, the lags model is found by information criteria AIC (see Table 6).

In ARMA models, three criteria set an optimal lag of 1. VAR models swing from 5 to 10 lags at maximum. SETAR, LSTAR, and STAR models presented an optimal lag length of 2 months. There is a concern in the long-term joint dynamics of these models, tested by the Johansen co-integration test (see Table 7).

Table 6. Optimal Lag Length - AIC

Model	Lag Length	AIC
Random Walk	Not Applied	2112.26
AR(p)	1	2100.15
MA(q)	1	2102.18
ARIMA(p,d,q)	1,1,0	2100.15
VAR A	9	11.62
VAR B	10	33.13
VAR C	5	73.90
VAR D	5	66.02
VAR E	5	69.30
VAR F	5	69.88
SETAR	2	1317.00
LSTAR	2	1320.00
STAR	2	1318.00

Source: Authors' elaboration.

The results in Table 7, which present Eigen and trace tests, are above their critical at 1%, which means all the models are cointegrated. Therefore, is desirable to treat these models as Vector Error Correction (VEC) by inserting an innovation in the main equation.

Table 7. Johansen Co-integration Test

	A	B	C	D	E	F
Statistic	$r = 0$					
λ_{eigen} (calculated)	27.53	85.32	78.78	101.02	101.02	104.02
λ_{eigen} (critical at 1%)	19.19	25.75	32.14	38.78	38.78	28.71
λ_{trace} (calculated)	30.21	95.96	105.21	151.44	101.02	93.18
λ_{trace} (critical at 1%)	23.52	37.22	55.43	78.87	38.78	34.24

Source: Authors' elaboration.

Finally, these models are estimated and their residuals tested to autocorrelation and heteroskedasticity, as shown in Table 8.

Table 8. Residuals Tests

Model	Lag Length	AIC
	<i>p-value</i>	<i>p-value</i>
Random Walk	0.1276	0.1200
AR(p)	0.1096	0.5500
MA(q)	0.0512	0.2012
ARIMA(p,d,q)	0.0913	0.0740
VAR A	0.1061	0.4586
VAR B	0.0941	0.1176
VAR C	0.1314	0.1683
VAR D	0.1121	0.9900
VAR E	0.1897	0.8639
SETAR	0.2014	0.1020
LSTAR	0.2812	0.5600
STAR	0.2412	0.1450

Source: Authors' elaboration.

Note: In univariate case, is used the test proposed in Box & Pierce (1970) and White (1980) in order to detect autocorrelation and heteroscedasticity.

As shown in Table 8, most models presented stability in their residuals. Therefore, the estimators of these models are consistent.

4. Results

In this section, we present the main results of this study. In section 4.1 is evaluated the forecasting accuracy in-sample and out-of-sample. The outcomes reveal to policy-makers, traders, agriculture and equity research analyst, hedgers, and portfolio managers, a deeper understanding of the best econometric model to forecast corn prices.

4.1 Forecasting Performance

The main signal of forecasting performance comes from the study of the loss function. In Table 9, the Root Mean Square Error measure is calculated to in-sample prediction.

Table 9. Model Accuracy Comparison

Model	RMSE
Random Walk	0.2098
AR(1)	0.0637
MA(7)	0.0676
ARIMA(2,1,1)	0.0582
VEC A	0.0522
VEC B	0.0527
VEC C	0.0564
VEC D	0.0552
VEC E	0.0567
SETAR	0.0570
STAR	0.0578
LSTAR	0.0571

Source: Authors' elaboration.

Table 9 is showed the RMSE results of each model. The lower in-sample RMSE is presented by model VEC A, the one that uses the price discovery role of futures contracts in the prediction. This model is the same one used by Xu (2018) for forecasting corn prices in the U.S. market. The results of the out-of-sample are shown in Table 10.

Table 10. Root Mean Square Error (10 months out-of-sample)

		1	2	3	4	5	6	7	8	9	10
RMSE 10 periods out-of-sample	Random Walk	0.0534	0.0987	0.1483	0.1924	0.2308	0.2200	0.2340	0.2277	0.2277	0.2342
	AR(p)	0.0981	0.1585	0.2183	0.2706	0.3154	0.3118	0.3314	0.3306	0.3354	0.3450
	MA(q)	0.0998	0.1592	0.2190	0.2707	0.3162	0.3133	0.3334	0.3332	0.3384	0.3494
	ARIMA (p,d,q)	0.1006	0.1640	0.2276	0.2732	0.3196	0.3160	0.3346	0.3368	0.3436	0.3556
	VEC A	0.0840	0.1197	0.1535	0.1698	0.1824	0.1449	0.1362	0.1077	0.0877	0.0765
	VEC B	0.0676	0.0894	0.1128	0.1143	0.1092	0.0474	0.0145	0.0337	0.0636	0.0779
	VEC C	0.0776	0.1106	0.1648	0.1999	0.2345	0.2190	0.2356	0.2349	0.2363	0.2487
	VEC D	0.0422	0.0481	0.0758	0.0835	0.0894	0.0349	0.0110	0.0372	0.0822	0.1158
	VEC E	0.0923	0.1334	0.1759	0.2043	0.2229	0.1877	0.1815	0.1579	0.1466	0.1473
	VEC F	0.0541	0.0339	0.0469	0.0195	0.0138	0.0447	0.0147	0.0267	0.0093	0.0134
	SETAR	0.0807	0.0674	0.0643	0.0533	0.0427	0.0137	0.0049	0.0219	0.0201	0.0169
	STAR	0.1118	0.0960	0.1026	0.0346	0.0460	0.0104	0.0112	0.0112	0.0189	0.0210
LSTAR	0.0583	0.0574	0.1254	0.0570	0.0653	0.0097	0.0288	0.0064	0.0020	0.0480	

Source: Authors' elaboration.

Despite the above result, we can only affirm that VEC F has better out-of-sample performance if it survives the Diebold-Mariano pairing method for the loss function prediction errors. In that regard, the Diebold-Mariano pairing method is useful to test the statistical difference among forecasts see Table 11.

Table 11. Diebold-Mariano Test

Model	Random Walk	AR(5)	MA(2)	ARIMA (11,1,2)	VEC F	SETAR	STAR	LSTAR
Random Walk	-	0.2022	0.1618	0.0347	0.0620	0.0133	0.0243	0.0089
AR(5)	0.7978	-	0.4416	0.0532	0.0623	0.0217	0.0349	0.0118
MA(2)	0.8382	0.5584	-	0.1610	0.1308	0.0081	0.0440	0.0021
ARIMA (11,1,2)	0.9652	0.9558	0.8689	-	0.2893	0.1372	0.0323	0.0271
VEC F	0.9380	0.9377	0.6116	0.7107	-	0.3414	0.3819	0.2174
SETAR	0.9867	0.9783	0.9991	0.8628	0.6586	-	0.4967	0.5103
STAR	0.9757	0.9651	0.8620	0.9677	0.6181	0.5033	-	0.4946
LSTAR	0.9911	0.9882	0.9967	0.9729	0.7826	0.4897	0.5054	-

Source: Authors' elaboration.

Note: the higher the probabilities, the better the model is.

Given the results in Table 11, the SETAR model has greater performance than other models regarding out-of-sample forecasting. In second place is the LSTAR. When considering the loss function of forecasting errors, the Diebold-Mariano test points out the SETAR and LSTAR model as more accurate than the VEC F model. This result is very interesting once it shows that even when including corn market fundamentals and macroeconomics variables, univariate models that treated structural breaks and non-linearities performed better. This finding corroborates the importance of studying the dynamics of corn prices, as previously done by other authors regarding other commodities, such as crude oil, by Albuquerque et al. (2018). In this sense, corn price series are similar to oil price series, as world demand and real economic activity can suddenly influence its long-run trend, causing a structural break.

5. Robustness test

This section provides a robustness test for forecasting corn prices. The robustness exercise used is the exclusion of the first difference in the variables.

Table 12. Performance comparison of models.

Model	RMSE
VEC A	0.0498
VEC B	0.0603
VEC C	0.0589
VEC D	0.0573

Source: Authors' elaboration.

In our estimates, the main models for forecasting corn prices within the sample were VEC A, VEC B, VEC C, and VEC D, with VEC A having greater predictive power, as shown in Table 9. Our robustness test corroborates the importance of including the future price in a vector autoregression since even excluding the first difference in the indicators, VEC A remains the main model for forecasting corn prices. This result is shown in Table 12.

6. Conclusion

Given the relevance of corn for food and fuel industries, scholars and financial market practitioners are constantly testing new models for corn price forecasting. This latter has the

purpose of achieving the highest forecasting accuracy, reducing trading costs, and enhancing hedging practices against volatility in prices. This paper deepens in this topic, estimating and comparing numerous econometric models.

Once it is common the presence of several shifts and level changes in time series, we have decided to investigate it in corn prices. The empirical exercise found a statistically significant structural break in 2010. Then, we have checked if the treatment of these problems could enhance the forecasting performance. For this purpose, we used Self-Exciting Threshold Autoregressive (SETAR) family models.

The estimation and comparison of SETAR family models and models commonly employed in literature, such as ARIMA and VEC, provided useful findings. The main results point out that the SETAR model performed better than others for out-of-sample forecasting. Considering in-sample forecasting, the best models were VEC A, having the lowest RMSE. These results highlight the importance of the previous treatment of structural breaks and non-linearities regarding corn price series, as it helps in achieving higher performance.

The results acknowledge market practitioners and researchers about important facts, such: i) the data treatment considering structural breaks and non-linearities helps in forecasting corn prices; ii) models that accommodate structural breaks can easily be implemented for forecasting, speculation, and hedge purposes, in this sense we strongly recommend the use of SETAR family models for worldwide corn market analysis.

The present work was able to achieve the objectives and highlight which are the most suitable models to be used in maize forecasting. However, the work has some limitations. A relevant limitation is that the study did not use models that are capable of dealing with a large number of predictors, that is, large (sparse) models. Among this class of models, the factor models, machine learning and deep learning models deserve to be highlighted. Currently, due to the high availability of data, high dimension models are gaining prominence in the forecasting exercise, as the lecture indicates that such a class of models may be more accurate than traditional models. And the models we use in the study have the limitation of being able to deal with a small set of predictors. Also, the variable of interest can be influenced by high-frequency (daily) variables, and the work left the Midas model out of the forecast model mix. The latter can estimate a dependent variable with a frequency different from its predictors.

Based on the limitations mentioned above, it is suggested to carry out forecasting exercises that include a larger number of predictors, for the use of high-dimension models. Also, it is important to investigate the use of the Midas model.

References

- Ahumada, H., & Cornejo, M. (2016). Forecasting food prices: the case of corn, soybeans and wheat. *International Journal of Forecasting*, 32(3), 838-848.
- Albuquerque, V. P., Medeiros, R. K., Besarria, C. N., & Maia, S. F. (2018). Forecasting crude oil price: does exist an optimal econometric model? *Energy*, 155, 578-591.
- Alvim, A. M., & Waquil, P. D. (2005). Efeitos do acordo entre o Mercosul e a União Européia sobre os mercados de grãos. *Revista de Economia e Sociologia Rural*, 43(4), 703-723.
- Baffes, J., & Haniotis, T. (2010). *Placing the 2006/08 commodity price boom into perspective*. Washington, DC: World Bank.
- Bastianin, A., Galeotti, M., & Manera, M. (2014). Causality and predictability in distribution: the ethanol-food price relation revisited. *Energy Economics*, 42, 152-160.

- Beckers, B., & Beidas-Strom, S. (2015). Forecasting the Nominal Brent Oil Price with VARs—one model fits all? Washington: International Monetary Fund.
- Benavides, G. (2009). Price volatility forecasts for agricultural commodities: an application of volatility models, option implieds and composite approaches for futures prices of corn and wheat. *Economics*, 3(2), 40-59.
- Bobenrieth, H. E. S., Bobenrieth, H. J. R., & Wright, B. D. (2004). A model of supply of storage. *Economic Development and Cultural Change*, 52(3), 605-616.
- Box, G. E., & Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, 65(332), 1509-1526.
- Campos, B. C. (2020). Are there asymmetric relations between real interest rates and agricultural commodity prices? Testing for threshold effects of US real interest rates and adjusted wheat, corn, and soybean prices. *Empirical Economics*, 59(1), 371-394.
- Chan, K. S., & Tong, H. (1986). On estimating thresholds in autoregressive models. *Journal of Time Series Analysis*, 7(3), 179-190.
- Cuaresma, J.C., Hlouskova, J., & Obersteiner, M. (2018). Fundamentals, speculation or macroeconomic conditions? Modelling and forecasting Arabica coffee prices. *European Review of Agriculture Economics*, 45(4), 583-615.
- Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 20(1), 134-144.
- Enders, W. (2008). *Applied econometric time series*. Chichester: John Wiley & Sons.
- Favro, J., Caldarelli, C. E., & Camara, M. R. G. D. (2015). Modelo de análise da oferta de exportação de milho brasileira: 2001 a 2012. *Revista de Economia e Sociologia Rural*, 53(3), 455-476.
- Food and Agriculture Organization of the United Nations – FAO. (2018). *The future of food and agriculture. Alternative pathways to 2050*. Rome: FAO. Retrieved in 2019, April 15, from <http://www.fao.org/3/I8429EN/i8429en.pdf>
- Gallagher, P. (1986). US corn yield capacity and probability: estimation and forecasting with nonsymmetric disturbances. *North Central Journal of Agricultural Economics*, 8(1), 109-122.
- Gilbert, C. L. (2010). How to understand high food prices. *Journal of Agricultural Economics*, 61(2), 398-425.
- Gurge, A. C. (2011). Impactos da política americana de estímulo aos biocombustíveis sobre a produção agropecuária e o uso da terra. *Revista de Economia e Sociologia Rural*, 49(1), 181-213.
- Hoffman, L. A., Etienne, X. L., Irwin, S. H., Colino, E. V., & Toasa, J. I. (2015). Forecast performance of WASDE price projections for US corn. *Agricultural Economics*, 46(S1), 157-171.
- Hotelling, H. (1931). The economics of exhaustible resources. *Journal of Political Economy*, 39(2), 137-175.
- Jadhav, V., Chinnappa, R. B., & Gaddi, G. M. (2017). Application of ARIMA model for forecasting agricultural prices. *Journal of Agricultural Science and Technology*, 19(4), 981-992.
- Jayne, T. S., & Rashid, S. (2010). *The value of accurate crop production forecasts*. Michigan: Michigan State University.
- Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics & Control*, 12(2-3), 231-254.
- Krishnamurthy, A., & Vissing-Jorgensen, A. (2011). *The effects of quantitative easing on interest rates: channels and implications for policy*. Cambridge: National Bureau of Economic.

- Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Berlin: Springer Science & Business Media.
- Mallory, M. L., Irwin, S. H., & Hayes, D. J. (2012). How market efficiency and the theory of storage link corn and ethanol markets. *Energy Economics*, *34*(6), 2157-2166.
- Mariscal, R., & Powell, A. (2014). *Commodity price booms and breaks: detection, magnitude and implications for developing countries*. Washington, DC: Inter-American Development Bank.
- McPhail, L. L., Du, X., & Muhammad, A. (2012). Disentangling corn price volatility: the role of global demand, speculation, and energy. *Journal of Agricultural and Applied Economics*, *44*, 401-410.
- Osathanunkul, R., Khiewngamdee, C., Yamaka, W., & Sriboonchitta, S. (2018). The role of oil price in the forecasts of agricultural commodity prices. In *International Conference of the Thailand Econometrics Society*. Berlin: Springer.
- Reichsfeld, D. A., & Roache, S. K. (2011). *Do commodity futures help forecast spot prices?* Washington, DC: International Monetary Fund.
- Roberts, M. J., & Schlenker, W. (2009). World supply and demand of food commodity calories. *American Journal of Agricultural Economics*, *91*(5), 1235-1242.
- Roberts, M. J., & Schlenker, W. (2013). Identifying supply and demand elasticities of agricultural commodities: implications for the US ethanol mandate. *The American Economic Review*, *103*(6), 2265-2295.
- Runge, C. F., & Senauer, B. (2007). How biofuels could starve the poor. *Foreign Affairs*, *86*, 41.
- Serra, T., & Gil, J. M. (2013). Price volatility in food markets: can stock building mitigate price fluctuations? *European Review of Agriculture Economics*, *40*(3), 507-528.
- Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, *48*(1), 1-48.
- Trujillo-Barrera, A., Mallory, M., & Garcia, P. (2012). Volatility spillovers in US crude oil, ethanol, and corn futures markets. *Journal of Agricultural and Resource Economics*, *37*(2), 247-262.
- United States Department of Agriculture – USDA. (2015). *Agricultural Baseline Database: Corn*. Retrieved in 2019, March 15, from <https://data.ers.usda.gov/reports.aspx?ID=17821>
- United States Department of Agriculture – USDA. (2019). *Feed Outlook*. Retrieved in 2019, March 15, from <https://downloads.usda.library.cornell.edu/usda-esmis/files/44558d29f/3f462b892/m039kb803/FDS-19b.pdf>
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, *48*(4), 817-838.
- Winkelried, D. (2016). Piecewise linear trends and cycles in primary commodity prices. *Journal of International Money and Finance*, *64*, 196-213.
- Winkelried, D. (2018). Unit roots, flexible trends, and the Prebisch-Singer hypothesis. *Journal of Development Economics*, *132*, 1-17.
- Xu, X. (2018). Cointegration and price discovery in US corn cash and futures markets. *Empirical Economics*, *55*(4), 1889-1923.
- Zhang, D., Zang, G., Li, J., Ma, K., & Liu, H. (2018). Prediction of soybean price in China using QR-RBF neural network model. *Computers and Electronics in Agriculture*, *154*, 10-17.
- Zivot, E., & Andrews, D. W. K. (2002). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business & Economic Statistics*, *20*(1), 25-44.

Received: April 21, 2020.

Accepted: February 27, 2021.

JEL Classification: C53; Q41; Q43.