Pontifícia Universidade Católica do Rio de Janeiro

Rafael Baptista Palazzi

Essays on commodity markets. A nonlinear approach to understanding the price and the market behavior.

Tese de Doutorado

Thesis presented to the Programa de Pós- Graduação em Administração de Empresas of the Departamento de Administração, PUC-Rio as partial fullfilment of the requirements for the degree of Doutor em Administração de Empresas.

Advisor: Prof. Antonio Carlos Figueiredo Pinto

Rio de Janeiro

April 2022

Rafael Baptista Palazzi

Essays on commodity markets. A nonlinear approach to understanding the price and the market behavior.

Thesis presented to the Programa de Pós-graduação em Administração de Empresas of PUC-Rio in partial fulfillment of the requirements for the degree of Doutor em Ciências – Administração de Empresas. Approved by the Examination Committee.

Prof. Antonio Carlos Figueiredo Pinto
Orientador
Departamento de Administração – PUC-Rio

Prof. Marcelo Cabus Klotzle
Departamento de Administração – PUC-Rio

Prof. Waldemar Antonio da Rocha de Souza
Universidade Federal de Alagoas

Prof. Ata Assaf
Faculty of Management and Business University of Balamand

Prof. Erick Meira
Financiadora de Estudos e Projetos (FINEP)

Rio de Janeiro, April 1, 202

Rafael Baptista Palazzi

Graduated in Economics at Mackenzie Presbyterian University in 2007 and obtained her M.Sc. Degree in Economics from the Pontifical Catholic University in Sao Paulo in 2011.

Bibliographic data

Palazzi, Rafael Baptista

Essays on commodity markets : a nonlinear approach to understanding the price and the market behavior / Rafael Baptista Palazzi ; advisor: Antonio Carlos Figueiredo Pinto. – 2022.

105 f. : il. color. ; 30 cm

Tese (doutorado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Administração, 2022.

Inclui bibliografia

1. Administração – Teses. 2. Mercado de commodities. 3. Mercado de energia. 4. Modelos não lineares. 5. Especulação. 6. Financialização. I. Pinto, Antonio Carlos Figueiredo. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Administração. III. Título.

CDD: 658

Acknowledgments

I am grateful to my advisor Prof. Antonio Carlos Figueiredo, for his support during the past four years. I am also thankful to Prof. Marcelo Klotzle for his encouragement and guidance during this research. Prof. Waldemar Souza, whom I met in 2016 at an event at B3, encouraged me to apply for the Ph.D. at PUC-Rio. Since then, Prof. Waldemar has provided helpful advice at different stages of this research.

For providing feedback on several aspects of this research, I would like to thank Prof. Sophie van Huellen (University of Manchester) and Dr. Erick Meira. Prof. Scott Irwin and Prof. Michel Robe have hosted me at the Dept. of Agricultural and Consumer Economics during my fellowship at the University of Illinois Urbana-Champaign (UIUC). Also, I thank Prof. Brian Wright for hosting me at the Dept. of Agricultural Economics at the University of California – Berkeley between December 2018 and March 2019.

My colleagues, Arnaldo Nascimento, Gerson Junior and Marcio Pimenta, shared the challenges of the Ph.D. process and were valuable friends.

The Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) funded my fellowship to UIUC (#200935/2020-8). This study was financed in part by the Coordenação de Aperfeiçoaento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

To the secretaries Teresa and Gisele for their support.

To my wife, whose love, dedication, support and encouragement gave me the strength and confidence to conclude this research.

To my parents for all the education and moral values that made me who I am. My father-in-law and mother-in-law for their invaluable support and assistance during my stay in Rio de Janeiro. My family, siblings, in-laws, nephews, and nieces shared joyful moments that made this process easier.

In memory of my beloved grandmother Celina and my dog Brownie Zegna.

Abstract

Palazzi, Rafael Baptista; Figueiredo Pinto, Antonio Carlos (Advisor). **Essays on commodity markets. A nonlinear approach to understanding the price and the market behavior**. Rio de Janeiro, 2022. 105p. Tese de Doutorado – Departamento de Administração, Pontifícia Universidade Católica do Rio de Janeiro.

Commodity markets have become a new investment alternative for portfolio investors over the last fifteen years, in a process known as the financialization of commodity markets. Several studies have explained the reasons for this phenomenon (e.g., speculation and increase in biofuels' production), leading to a question largely understudied in agricultural and energy economics literature. How has the financialization of the commodities market changed the price dynamic over the years? This thesis applies nonlinear models to understand whether the speculation caused the price movements in the agricultural commodity markets; investigates the price discovery in the Brazilian market by analyzing the transmission of international energy and feedstocks prices to Brazilian ethanol and gasoline prices; and investigates the spillover effects from global futures markets to local spot prices. In addition, it analyzes the increased liquidity in the commodity markets by developing a new measurement to gauge the degree of ambiguity for 12 agricultural commodities prices. Despite the robust econometric tests performed, the findings were inconclusive on the role of speculation in impacting the price returns of commodities. It also found that there exists a nexus between international oil and Brazilian ethanol prices, and global commodities prices have increased the spillover effects on the Brazilian spot markets. Finally, the financialization of commodity markets has increased the liquidity in the market as measured by the degree of ambiguity. This thesis contributes to the field not only by applying more robust, novel econometric approaches but also by evidencing how information discovery and risk-sharing affect the commodity price dynamics.

Keywords

Commodity markets; nonlinear model; financialization; energy commodity; speculation; spillover effects; ambiguity.

Resumo

Palazzi, Rafael Baptista; Figueiredo Pinto, Antonio Carlos. **Ensaios sobre o mercado de commodities. Uma abordagem não linear para entender a dinâmica do preço e o comportamento do mercado**. Rio de Janeiro, 2022. 105p. Tese de Doutorado – Departamento de Administração, Pontifícia Universidade Católica do Rio de Janeiro.

Os mercados de commodities tornaram-se uma nova alternativa para investidores nos últimos quinze anos, em um processo conhecido como financeirização dos mercados de commodities. Vários estudos têm explicado as razões deste fenômeno, porém esta é uma questão ainda pouco estudada na literatura de economia agrícola e energética no Brasil. Como a financeirização do mercado de commodities mudou a dinâmica dos preços ao longo dos anos? Esta tese aplica modelos não lineares para entender se a especulação causou os movimentos de preços nos mercados de commodities agrícolas, bem como para investigar a descoberta de preços no mercado brasileiro ao se testar os mecanismos de transmissão dos preços internacionais de energia e commdities agrícolas aos preços brasileiros de etanol e gasolina. Procuramos investigar com os mesmos modelos não lineares os efeitos de transbordamento dos mercados globais de futuros para os preços à vista locais. Por fim, analisa-se o aumento da liquidez nos mercados de commodities, desenvolvemos para tanto uma nova medida para compreender o grau de ambiguidade dos preços de 12 commodities agrícolas. Apesar dos testes econométricos, os resultados foram inconclusivos sobre o papel da especulação no impacto dos retornos dos preços das commodities. Existe um nexo entre os preços internacionais do petróleo e do etanol brasileiro, e os preços globais das commodities aumentaram os efeitos de contágio nos mercados spot brasileiros. Finalmente, a financeirização dos mercados de commodities aumentou a liquidez do mercado medida pelo grau de ambiguidade. Esta tese contribui para o campo ao aplicar abordagens econométricas robustas e inovadoras, bem como ao evidenciar como o *price discovery* e o *risk-sharing* afetam a dinâmica dos preços das commodities.

Palavras-chave

Mercado de commodities; mercado de energia; modelos não lineares; especulação; financialização; contágio; ambiguidade.

Table of Contents

1 Introduction 7

2 Can We Still Blame Index Funds For The Price Movements In The Agricultural Commodities Market? 9
2.1 Introduction 9
2.2 Literature Review 10
2.3 Data 12
2.4 Methodology 20
2.5 Results 22
2.6 Conclusion 26

3 The Sugar-Ethanol-Oil Nexus In Brazil: Exploring The Pass-Through Of International Commodity Prices To National Fuel Prices 27
3.1 Introduction 27
3.2 Background 29
3.3 Data And Methodology 36
3.4 Results And Discussions 45
3.5 Conclusions, Implications And Suggestions For Future Studies 48
3.6 Appendix. Bivariate Nardl Models: Cumulative Dynamic Multipliers 49
3.7 Appendix. Rolling Regressions: Robustness Check 50

4 Volatility Spillovers In The Brazilian Agricultural Markets. 52
4.1 Introduction 52
4.2 Literature Review 52
4.3 Methodology And Data 54
4.4 Empirical Results And Discussion 56
4.5 Conclusion And Further Research 60
4.6 Appendix 61

5 Can We Measure Ambiguity In Commodity Futures Markets? Formulation Of An Empirical Ambiguity Measure 63
5.1 Introduction 63
5.2 Literature Review 63
5.3 Methods And Data 66
5.4 Results And Discussion 73
5.5 Summary And Conclusions 79
5.6 Appendix 81

6 Conclusion **86**

References **87**

1
Introduction

Commodity markets have become a new investment alternative for portfolio investors over the last fifteen years. A large inflow of investments from commodity index traders, hedge funds and other investors has entered the commodity markets seeking an alternative asset class. The amount of commodity index-related investment purchases increased from around $15 billion in 2003 to roughly $200 billion in 2008 (Tang & Xiong, 2012). This process was named the financialization of the commodities markets and caused the prices of non-energy commodity futures to become correlated with oil prices.

Nonetheless, trading activities that led food prices to soar called the attention of the US Congress and the US Commodity Futures Trading Commission (CFTC). The "Master Hypothesis" was named to address the long-only index investments as the main cause of the spike in commodity prices (Irwin, 2013). Scholarly literature investigated whether speculation impacted the level of commodity futures prices. For instance, Irwin and Sanders (2011) analyzed studies that evidenced the effects of index funds on commodity prices, criticizing the data and the methods applied. Grosche (2014) also emphasized the limitations of the stand-alone application of Granger Causality (GC) tests.

My thesis aims to investigate the financialization of the commodities market in Brazil, a research agenda that is yet to be developed in the country. Brazil ranks among the largest commodities producers and exporters in the world. Hence, understanding how the commodity market in Brazil has become central to speculation and its underlying effects is crucial for designing better regulatory and pricing policies, as well as for the proposal of optimal hedging strategies. From an academic perspective, my study contributes to the discourse by developing new approaches to assess the process of financialization and by proposing a new metric to gauge ambiguity in the commodities markets. My study also analyzes the symmetrical and asymmetrical effects of the transmission mechanisms from international futures prices to local markets in a multivariate framework and the spillover effects of the volatility from global futures prices to spot prices in the Brazilian market. This thesis is organized into four chapters.

The first chapter, which was published as an article in the *International Journal of Economics and Finance*, deals with the controversial role of speculation in the agricultural commodity markets. It innovates by using a robust econometric approach to test whether speculation impacted the price returns of 12 agricultural commodities from 2006 to 2018. In contrast to previous studies that considered only linear methods to test the impact of speculation in the commodities prices/returns or volatility, I proposed to compare linear and nonlinear methodologies not only in bullish but also in bearish markets. The main results were inconclusive as to whether speculation and index funds influenced the returns of the agricultural commodity prices. Even my robust approach could not detect an effect of speculation on the prices' return.

Besides the financialization investigated in the first chapter, other factors may have led to the surge in commodities prices. For example, biofuel production is one of the main causes of rising food prices (Mitchell, 2008). Thus, my second chapter analyzes the pass-through of international commodity prices to the local fuel markets. An experiment is conducted in the Brazilian context. This chapter provides a first insight into the asymmetric effects of global energy and agricultural commodity prices on fuel prices in Brazil from 2015 to 2020. Multivariate nonlinear autoregressive distributed lag (NARDL) models are built using weekly prices of energy commodities traded in futures markets. Heating oil (HO), reformulated blendstock for oxygenate blending (RBOB), West Texas Intermediate (WTI) crude oil, Chicago ethanol (Platts), corn and sugar are considered as independent variables in the models, whilst the Brazilian ethanol and domestic gasoline prices represent the dependent variables. Previous studies have overlooked the interaction between international energy, feedstock prices and domestic fuel prices in a multivariate framework. Moreover, few studies have focused on the pass-through mechanisms of global commodity prices to domestic ethanol or gasoline prices. Our findings suggest that prices for the RBOB have had both short- and long-run asymmetric effects on domestic gasoline prices in Brazil, while surges in HO prices have led to declines in Brazilian ethanol prices.

The third chapter expands the analysis of the interconnection between international and spot prices in the Brazilian context while remaining focused on the spillover effects, assessing the transmission of futures prices traded at the global futures markets and spot prices in Brazil. It aims to investigate the volatility spillovers among coffee, ethanol, soybeans, RBOB futures prices, and Brazilian spot prices from 2010 to 2020, using the Diebold and Yilmaz (DY) volatility spillover analytical framework to estimate the total, the gross and the net directional volatility spillover. Our results demonstrated an increasing trend in the total volatility spillover index, suggesting an increase in the Brazilian market's connectedness with international markets.

Finally, the financialization of the commodities has contributed to the liquidity in the market, mainly by appealing to the noncommercial traders that seek an alternative investment/asset class. Thus, the last chapter proposes to estimate ambiguity in commodity futures markets by formulating an empirical ambiguity measure. We compose an empirical measure to explain the 12 commodity prices of the nearest futures contracts. Ambiguity is defined as uncertainty in the probability distribution due to misinterpretation or lack of information and is a current feature of financial assets. The results show that ambiguity represents a current feature of commodity futures prices that demonstrates a significant autoregressive pattern coupled with weak and mixed impacts in the log returns.

2
Can we still blame index funds for the price movements in the agricultural commodities market?

2.1
Introduction

From the early 2000s onwards, the commodities market underwent a profound transformation. Many participants entered the market to diversify their investments. Irwin and Sanders (2011) point out that more than US$ 100 billion of new resources entered this market from 2004 to 2008, a process that became known as the financialization of the commodities market. Coincidentally, commodity prices exploded. Cheng and Xiong (2014) show that the WTI price reached US$ 150 per barrel by 2008. Likewise, food prices also increased, a phenomena that had been not observed since the 1970s (Carter et al., 2011; FAO, 2008). Based on these trends, a series of studies by academics and international agencies examined the role of speculation in raising commodity prices (International Monetary Fund, 2008; Irwin and Sanders, 2011, et al., 2009). Index funds were the first source blamed for the price distortion, and international organizations were publicly exposed. For example, Irwin et al. (2011) cite a report from the US Senate (2014) subcommittee pointing to the Commodity Index Traders (CIT) as the responsible for raising prices on wheat futures contracts.

Some researchers found it difficult to establish a causality between the position of speculative funds and the level of commodity prices. Many of these studies used the Granger Causality (GC) test. However, this test has several limitations (E. M. de Oliveira et al., 2019). Hiemstra and Jones (1994) argue that, although the GC test has the power to show causality in linear relations, it does not offer the same efficiency in nonlinear relations. Grosche (2014) points out limitations in this type of analysis, arguing that a review of these studies demonstrates that the findings are inconclusive. Sanders & Irwin (2017) also identify some criticisms of GC tests, such as a low-power test in volatile commodity futures markets, efficient markets with restricted information, and lack of conditioning variables within models. Given the restrictions in current models, would models that overcome these limitations capture these relationships better?

In this study, we examine the influence of CIT positions and excessive speculation measured by the T index of speculation, following Working (1960) on agricultural commodity prices. We use weekly data from the Commodities Futures Trading Commission (CFTC) for the period between 2006 and 2018. We apply linear and nonlinear models of causality to determine whether the speculative performance in the agricultural commodities market generated changes in the prices of the commodities. We begin by considering the classical causality model (Granger, 1969) and using the first differences of the series. Next, we test the existence of causality using Toda and Yamamoto's (1995) methodology, which deals with the possibility of cointegration between the variables. Finally, we employ the Diks and Panchenko's (2005, 2006) nonlinear model, which tests for non-parametric causality. After running the tests, we use filters to control heteroscedasticity with multivariate GARCH models, specifically for nonlinear causality and the Granger BEKK-GARCH tests (Baba et al., 1990; Bollerslev et al., 1988; Engle & Kroner, 1993).

The paper proceeds as follows. Section 2 discusses the literature on the relationship between speculation and price changes. One group finds evidence of a relationship between speculation and prices, while another group finds no evidence of causality. Section 3 describes the data and the proxy for speculation. Section 4 presents the econometric methodology to test the causal relationships between the variables. Section 5 presents and discusses the model results. Section 6 concludes and offers suggestions for future research.

2.2
Literature Review

Prior research addresses the causal relationship between index funds and commodity prices. In this context, Irwin and Sanders (2012) consider the hypothesis that long-only index investments were the principal causal agents of a hike in future commodity prices as the master hypothesis. The authors test the hypothesis using the GC method but find no evidence to support this proposition. A significant part of the literature finds no evidence or a limited relationship (Aulerich et al., 2013; Bohl et al., 2012; Capelle-Blancard & Coulibaly, 2011; Gilbert, 2010; Gilbert & Pfuderer, 2014; Irwin, 2013; Irwin et al., 2016; Lehecka, 2015; Vercammen & Doroudian, 2014), thus providing evidence that is inconsistent with the master hypothesis.

Some studies examine this hypothesis in a similar way and employ the GC test with different approaches. For example, Aulerich et al. (2010) test the GC between the change in the long position of the CIT and the return on the futures contract of the closest screen. The authors test two distinct periods, from 2004 to 2005 and from 2006 to 2008, to allow the rolled positions to appear in the sample. However, they could not prove any relationship between speculation and price variation. Gilbert (2010) also finds no causality between speculation (referred to as non-commercial) and returns on commodity prices such as corn, soybeans, and oats. Gilbert and Pfuderer (2014) expand the role of agricultural commodities to include markets with lower liquidity such as the live cattle market and lean hogs. They reach a similar conclusion, even for markets with lower liquidity. Capelle-Blancard and Coulibaly (2011) innovated the test using a seemingly unrelated regression system, which permits a test of causality that accounts for the correlation between different markets. The result shows that, between 2006 and 2010, index funds were not responsible for the increase in agricultural commodity prices. Lehecka (2015) employs the causality test using the Toda and Yamamoto model for different measures of speculation and shows that the results are not very useful in explaining the behavior of prices in the face of speculative pressure, or pressure by hedgers.

Another stream of literature suggests that speculation does not cause the market bubble in agricultural commodities (Bohl et al., 2012; Irwin et al., 2009, 2016; Lancaster, 1989; Sanders & Irwin, 2011; Vercammen & Doroudian, 2014). On the contrary, the performance of the funds can be seen as beneficial to the financial market, as it generates liquidity and can reduce market volatility. Thus, Brunetti et al. (2016) argue that the increase in speculative positions in the commodities market is benign because investors can help reduce volatility by taking positions contrary to hedgers.

Another advantage Irwin et al. (2011) point out is that the expansion of speculative activity in the commodity market may reduce the risk premium and, consequently, the cost of hedging.

However, the literature is not unanimous on the benefits of speculative funds in the commodity market. Some researchers show that such funds may not only trigger an increase in commodity prices, but also increase volatility. Robles et al. (2009) use the spot prices of agricultural commodities (wheat, corn, rice, and soy) supplied by the Food and Agriculture Organization of the United Nations and five variables as proxies for speculation. The GC test shows that speculative activities may influence the prices of these commodities, which may have negative consequences in terms of dealing with hunger in underdeveloped countries. Gutierrez (2013) finds evidence that commodity prices deviated from their equilibrium price from 2007 to 2008. The author uses the bootstrap method to compute the sample probability distribution, whose alternative hypothesis is located in the right part of the distribution. The author then uses the unit root test, which determines the explosive part of the statistical test. The author finds that the first sign of the deviation in the wheat market was in August 2007, followed by corn shortly after, in February of 2008.

Although some authors reach different conclusions, some criticize the methods applied in these studies. Grosche (2014) exposes certain limitations: (i) the failure to consider the informational efficiency of the markets, the variation of time, and the response to the effects of the bounded rationality of the different trading strategies; and (ii) limitations in the CFTC's data, which they disclose only weekly, and causality, in the sense that price relative to speculative funds does not necessarily mean that funds follow prices (trend followers). The study recommends the use of multivariate models (vector autoregressive, or VAR) as well as nonlinear models of causality. In this sense, Sanders and Irwin (2017) also acknowledge the limitations of GC tests. To address criticisms of GC, the authors apply a time-series correlation between the Supplemental Commitments of Traders report (SCOT) and nearby futures returns, as well as a cross-section correlation between the Index Investment Data report (IID) and nearby futures returns. The results show a positive correlation between changes in SCOT and nearby futures returns, mainly in years that were not bubble-like. Moreover, shifts in SCOT index position did not coincide with changes in price. The authors also tested the correlation between SCOT index positions and daily market returns. Results show no correlation between index positions and daily market returns. Cross-sectional analysis of IID buy-positions and market returns fails to provide evidence of a relationship, even controlling for other macroeconomic variables.

Baek and Brock (1992) show that linear causality tests have little power to detect nonlinear relationships over time. Thus, some studies improve the analysis from non-parametric econometric tests. Hiemstra and Jones (1994) modify the Baek and Brock version by relaxing the hypothesis that the time series is independent and identically distributed (iid) to test the dynamic relationships. Diks and Panchenko (2006) refine the Hiemstra and Jones test to overcome the possibility of spurious rejection of the null hypothesis. Another limitation is the possibility of cointegration effects in the causality tests in linear and nonlinear series. Bekiros and Diks (2008a) investigate the relationship between the spot market and the future WTI market empirically and find that even after applying appropriate models in the BEKK-GARCH family to control conditional heteroscedasticity, the nonlinear relationship persists in some cases, while the linear relationship disappears after filtering.

Therefore, the hypothesis that speculation causes a price increase requires further investigation in light of the more sophisticated models presented in recent work.

2.3
Data

2.3.1
Commodities prices and the CIT position

We collected the futures prices of ten agricultural commodities: corn, cattle (feeder cattle and live cattle), lean hog, soybeans, and wheat from the Chicago Mercantile Exchange (CME), and coffee, cocoa, and sugar from ICE Futures, US, in New York. The prices refer to the contracts of the first active screen of the exchanges.

We use the CIT position for testing the Master Hypothesis. It is worth noting that depending on the purpose of the speculative hypothesis, a different measure should be applied. That is, (Etienne, Irwin, & Garcia, 2018) provide sharp distinctions between speculative measures, such as Index Investments, non-commercial activities, the Working's T index, and the ESV Index. According to the authors, the CIT net position is used to test the Master Hypothesis of whether the net long position can cause a spike in agricultural commodity prices. Working's T index reflects the excess of speculation relative to hedging movements, which could impact the agricultural futures prices.

The CFTC began releasing this report as a supplement to the Commitments of Traders (COT) in 2007 for only twelve agricultural commodities. In this supplement, CFTC removed the positions of the index traders from the non-commercial and commercial positions and separated them by category: CIT long and short. CFTC also removed managed funds, pension funds, and other institutional investors which seek exposure in commodity markets as an asset class, from the position of non-commercials. Over-the-counter (OTC) hedge funds, such as swap dealers, are drawn from the position of the commercials. (Sanders, Irwin, & Merrin, 2010) show that 85% of the index trader 's position came from the long commercial category of the COT report. The authors argue that most of the long-only index positions initially trade in the OTC markets, then re-enter into futures markets to hedge their exposures through swap dealers (commercial and investments banks).

Even with the improvement in the classification of traders in this new report (SCOT), the CFTC itself warns about possible failures in the classification procedure (CFTC, 2008). Another limitation in the publication is the data horizon, which starts only in 2007. Irwin (2013) points out that in the three years prior to 2007, massive fund positions received training. In this case, between 2004 and 2006, there was a rapid increase in the CIT, with index traders' long positions almost tripling in the corn and wheat market. The author also shows that the commodity boom in 2007–2008 may be only a coincidence since the greatest pressure on CIT positions was from 2004 to 2006.

Despite its limitations, the SCOT report provides a useful proxy for the position of index traders, as (Irwin & Sanders, 2012a) point out since the errors associated with the report are small and still reflect the behaviors of index traders in the future agricultural commodities market.

2.3.2 Speculative Proxies

Widely used in the literature, the T index proposed by (Working, 1960) measures the excess of speculation in the market. The measure captures the excess of speculative positions in relation to hedge positions. The index is calculated as

$$T = \begin{cases} 1 + \frac{SS}{HL+HS};\, if\ HS > HL, or \\ 1 + \frac{SL}{HL+HS};\ if\ HS < HL \end{cases} \tag{1}$$

where *SL* and *SS* are the *Long Position* and *Short Position*, respectively, while *HL* and *HS* are the positions *hedgers* bought and sold. The minimum value is 1, so the excess beyond 1 reflects the level of speculation that *hedgers* did not absorb.

We consider the long and short CIT positions and analyze each of them separately. The master hypothesis is that the pressure of purchased funds created a bubble in the commodity market and had its apex in 2006–2008 (the price of sugar had its bubble afterward, in 2010–2011). However, some commodities fell over time, as Figure 1 shows. Thus, we include the CIT short position to analyze whether there was pressure on the movement of commodities separately. Our sample period runs from the beginning of the data provided in the SCOT report (2006 and 2007) until 2018.

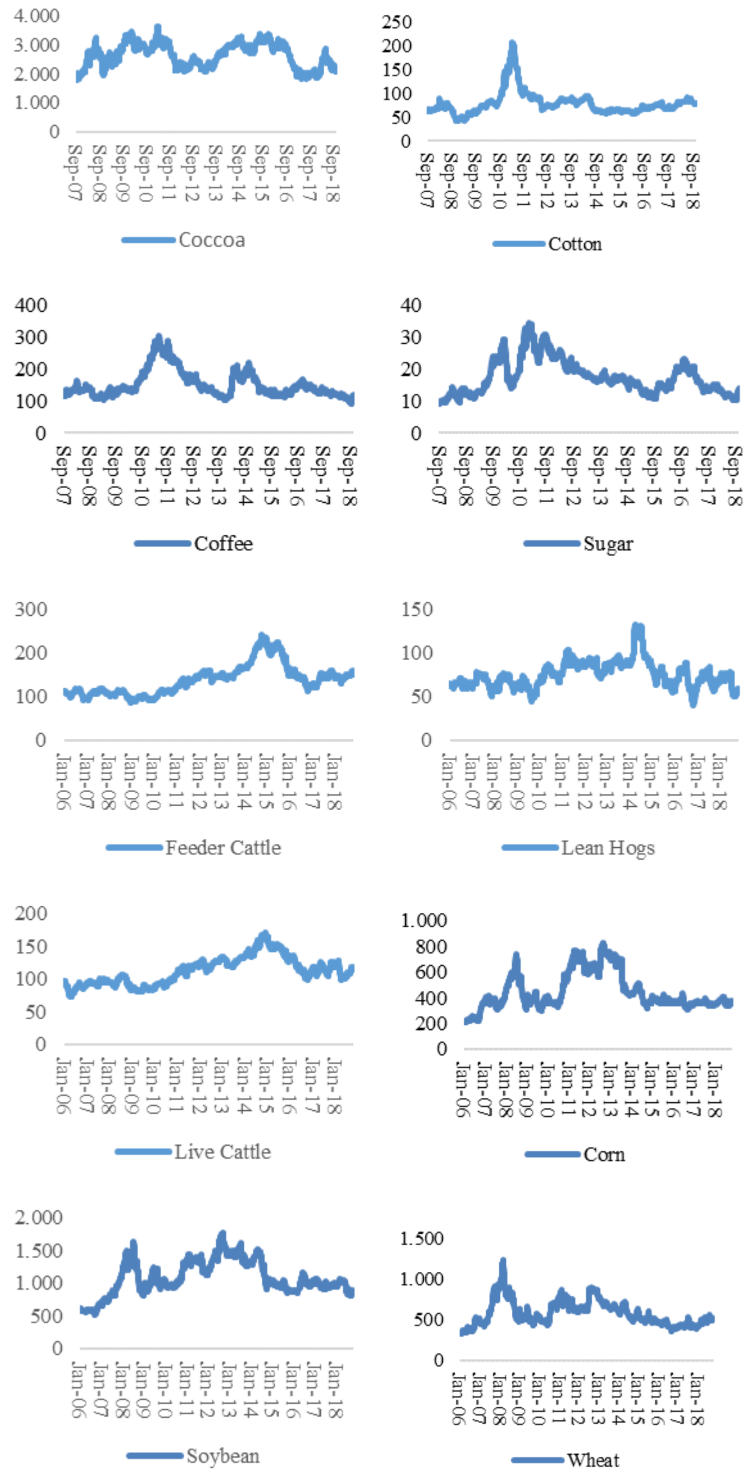


Figure 1. Commodity prices. The graphs show future commodity prices for the period between 2006–2007 and 2018. For each commodity, the contracts are measured on different scales: Cocoa – 10 metric tons; Cotton – 50,000 pounds; Coffee – 37,500 pounds; Sugar – 112,000 pounds; Feeder Cattle – 50,000 pounds; Lean Hogs – 40,000 pounds; Live Cattle – 40,000 pounds; Corn - 5,000 bushels; soybean – 5,000 bushels; wheat – 5,000 bushels. *Source: Bloomberg.*

Table 1 (panels A, B, C, and D) presents the descriptive statistics. The series are all non-stationary according to the Augmented Dickey–Fuller tests (measured by constant and constant plus intercept) for all data. In panel A of Table 1, the prices are transformed into log-returns through the following relationship: $r_i = \ln P_t - \ln P_{t-1}$. We see excess of kurtosis for the price series exhibiting heavy (leptokurtic) tails. We find negative asymmetry in more than half of the series. In Panel B, we analyze the T index, which shows no negative asymmetry for any commodity, though we find excess kurtosis in all series. To detect the null hypothesis of normality, we apply the Jarque–Bera test, and thus reject the null hypothesis of normality for all commodities, both prices and positions, using the speculation index. Panels C and D show the log of the return of the positions of funds bought and sold.

	Mean	Median	Max	Min	Std. Dev.	Skewness	Kurtosis	J-B	Obs	DF (c.)	DF (ct)
Panel A: Return on prices (P_t)											
Cocoa	0.000	0.002	0.154	-0.124	0.041	0.488	4.236	124.590***	1,206	-9.201***	-9.200***
Cotton	0.000	0.002	0.162	-0.145	0.040	-0.192	4.153	35.777***	581	-23.653***	-23.637***
Feeder Cattle	0.000	0.002	0.096	-0.120	0.024	-0.374	4.469	75.589***	668	-25.756***	-25.737***
Lean Hogs	-	-0.001	0.205	-0.224	0.047	0.231	6.975	445.795***	668	-25.708***	-25.706***
Live Cattle	0.000	0.001	0.093	-0.110	0.024	-0.192	4.542	70.296***	668	-27.769***	-27.751***
Coffee	0.000	-0.000	0.177	-0.145	0.044	0.082	3.663	11.294***	581	-24.279***	-24.278***
Corn	0.001	0.002	0.233	-0.256	0.044	-0.179	6.869	420.143***	668	-27.243***	-27.292***
Soybean	0.000	0.002	0.120	-0.200	0.035	-0.528	5.300	178.291***	668	-26.867***	-26.941***
Sugar	0.001	-0.001	0.171	-0.230	0.050	-0.101	4.370	46.431***	581	-25.850***	-25.889***
Wheat	0.001	- 0.002	0.169	-0.176	0.047	0.214	3.720	19.508***	668	-26.529***	-26.542***

Table 1. Descriptive statistics and unit root tests (price series)

Notes: Descriptive statistics on the weekly observations of 10 commodities presented between 2006–2007 and 2018. The log-returns of the series prices ($\boldsymbol{r_i = \ln P_t - \ln P_{t-1}}$) of commodities in the Chicago (CME) or New York (ICE Futures US) exchanges are source from the Bloomberg platform. *, **, and *** show significance at the 10%, 5%, and 1% level, respectively. Augmented Dickey–Fuller (DF) tests were applied, as measured by the constant (DF c.) and the (DF ct) intercept.

	Mean	Median	Max	Min	Std. Dev.	Skewness	Kurtosis	J-B	Obs	DF (c.)	DF (ct)
Panel B: Speculation index (T)											
Cocoa	1.197	1.161	2.048	1.017	0.139	2.222	9.793	3,311.146***	1,206	-5.778***	-6.375***
Cotton	1.111	1.070	1.569	1.011	0.097	1.762	6.786	648.722 ***	581	-4.660***	-4.715***
Feeder Cattle	1.421	1.396	2.729	1.052	0.215	1.484	7.711	864.163***	668	-6.159***	-6.314***
Lean Hogs	1.269	1.223	1.927	1.028	0.166	1.079	3.990	157.130***	668	-4.638***	-4.634***
Live Cattle	1.178	1.165	1.407	1.036	0.095	0.418	2.061	44.107***	669	--4.420***	-4.561***
Coffee	1.183	1.157	1.553	1.028	0.128	1.128	3.642	133.473***	581	-4.008***	-4.227***
Corn	1.167	1.138	1.453	1.032	0.110	0.816	2.562	79.624***	668	-3.972***	-5.293***
Soybean	1.131	1.105	1.472	1.034	0.080	1.308	4.736	274.903***	668	-4.718***	-4.761***
Sugar	1.135	1.082	1.430	1.005	0.117	0.733	2.129	70.467***	581	-3.047**	-3.857**
Wheat	1.419	1.379	1.993	1.091	0.202	0.602	2.389	50.803***	668	-3.866***	-4.191***

Notes: Descriptive statistics on the weekly observations of 10 commodities presented between 2006–2007 and 2018 in terms of the index of speculation (T). *, **, and *** show significance at the 10%, 5%, and 1% level, respectively. Augmented DF tests were applied, as measured by the constant (DF c.) and the (DF ct) intercept.

	Mean	Median	Max	Min	Std. Dev.	Skewness	Kurtosis	J-B	Obs	DF (c.)	DF (ct)
Panel C: Long position of the *index traders*											
Cocoa	0.001	0.004	2.155	-2.078	0.137	-1.597	157.966	1,206,244.168***	1205	-34.643***	-34.632***
Cotton	0.000	0.000	0.110	-0.122	0.028	-0.426	5.544	174.172***	581	-21.597***	-21.645***
Feeder Cattle	0.002	0.003	0.161	-0.157	0.043	-0.178	4.508	66.783***	668	-24.856***	-24.84877***
Lean Hogs	0.001	0.002	0.127	-0.265	0.028	-1.060	16.908	5,508.944***	668	-22.947***	-22.932***
Live Cattle	0.001	0.002	0.078	-0.116	0.021	-0.350	6.555	365.488***	668	-9.904***	-9.880***
Coffee	0.001	0.001	0.166	-0.260	0.033	-0.503	12.849	2,372.528***	581	-20.448***	-20.448***
Corn	0.001	0.002	0.154	-0.129	0.026	-0.167	8.000	699.013***	668	-23.659***	-23.657***
Soybean	0.001	0.003	0.139	-0.140	0.029	-0.582	7.116	509.242***	668	-22.523***	-22.514***
Sugar	0.000	0.001	0.101	-0.094	0.025	-0.138	5.248	124.230***	581	-20.017***	-19.999***
Wheat	-0.001	0.002	0.181	-0.843	0.057	-8.419	121.310	397,480.074***	668	-25.531***	-25.51953***

Notes: Descriptive statistics on the weekly observations of 10 commodities presented between 2006–2007 and 2018, in terms of the returns of the long positions of the *index traders* (CIT). *, **, and *** show significance at the 10%, 5%, and 1% level, respectively. Augmented DF tests were applied, as measured by the constant (DF c.) and the (DF ct) intercept.

	Mean	Median	Max	Min	Std. Dev	Skewness	Kurtosis	J-B	Obs	DF (c.)	DF (ct)
Painel D: Short position of the *index traders*											
Cocoa	0.004	0.014	18.975	-18.534	1.103	0.235	241.980	2,867,485.111***	1205	-21.582***	-21.575***
Cotton	0.003	0.020	1.773	-1.206	0.270	-0.718	9.841	1,182.794***	581	-19.868***	-19.855***
Feeder Cattle	0.002	0.000	18.980	-19.787	2.755	-0.386	37.240	32,647.828***	668	-6.304***	-6.323***
Lean Hogs	0.006	0.001	18.159	-16.524	1.529	0.872	94.043	230,790.975***	668	-13.045***	-13.120***
Live Cattle	0.003	0.014	16.013	-20.699	1.112	-5.113	245.346	1,637,601.370***	668	-20.928***	-20.912***
Coffee	0.007	0.018	1.438	-1.919	0.245	-0.795	16.572	4,520.386***	581	-20.787***	-20.770***
Corn	0.003	0.021	0.927	-1.630	0.182	-2.058	18.938	7,541.748***	668	-26.693***	-20.315***
Soybean	0.004	0.013	1.436	-0.952	0.189	-0.330	12.627	2,591.544***	668	-26.074***	-26.055***
Sugar	0.001	0.015	0.438	-0.708	0.136	-1.295	7.353	621.138***	581	-18.350***	-18.335***
Wheat	0.003	0.008	1.241	-1.687	0.199	-0.605	16.348	5,000.014***	668	-19.160***	-19.193***

Notes: Descriptive statistics on the weekly observations of 10 commodities presented between 2006–2007 and 2018. The returns on the sell positions of the index traders' (CIT) are presented. *, **, and *** show significance at the 10%, 5%, and 1% level, respectively. Augmented DF tests were applied, as measured by the constant (DF c.) and the (DF ct) intercept.

With current data and a broader window, we may observe that prices returned to the prevailing levels prior to the 2008 bubble. Therefore, an up-to-date analysis of the relationship between speculation and commodity prices is important in terms of the other studies of the topic.

2.4 Methodology

2.4.1 Granger Causality (GC)

GC, much used in the empirical literature, analyzes the dependency relationship between variables over time (Granger, 1980). Irwin (2013), who published several studies on this subject using this model, argues that it is already established that the results of GC tests should be interpreted with caution. For example, he cites studies that state that rejecting the null hypothesis of Granger's causation may not reflect a true causal relationship between x and y, but rather, the omission of variable z, which may be the cause of both x and y.

2.4.2 Augmented Granger-causality test

The GC test was modified over time to capture the dynamics of the cointegration relationship in time series using the VAR model. In this connection, Granger (1988) shows that, in some cases, the traditional GC test does not detect causal relationships when the series are cointegrated. Engel and Granger (1987) point out that if variables x and y are non-stationary and cointegrated, then this could invalidate the results of the GC test. To address the problems of cointegration and unit roots, (Sims, Stock, & W Watson, 1990) show that, in a system that includes a unit root, using a VAR model to test constraints on coefficients does not apply to different orders of integration. Thus, Toda and Yamamoto (1995) propose a solution that deals with the problem of stationarity and cointegration. Their model guarantees that the chi-square (X) distribution is asymptotic (modified Wald statistical test). The procedure is to determine the optimal *lag* (k) based on some criterion, such as (Akaike, 1974) Akaike (1974) or and Schwarz (1978) (Bayesian information criteria); the order (k + d_{max}) is estimated, where *d* is the maximum order of integration of the model

$$Y_t = \alpha + \sum_{i=1}^{h+d} \beta_i \, Y_{t-i} + \sum_{i=1}^{k+d} \gamma_j \, X_{t-j} + u_{yt} \qquad (2)$$

$$X_t = \alpha + \sum_{i=1}^{h+d} \theta_i \, Y_{t-i} + \sum_{i=1}^{k+d} \delta_j \, X_{t-j} + u_{xt}, \qquad (3)$$

where *h* and *k* are the *lag lengths* of Y_t and X_t; $u_t's$ are the assumed error terms as white noise. The estimation of VAR (*k* + *d*) makes it possible for the X^2 distribution of the Wald test with *k* degrees of freedom to be admitted. Thus, the *lag length* of the model should be smaller than the integration order.

2.4.3
Nonlinear causality — Diks and Panchenko (Diks & Panchenko, 2006)

Linear models of causality are more common in the literature. As outlined in the previous section, the traditional GC model has little power to detect the nonlinear behavior of variables (persistence or structural breaks, for example). Hiemstra & Jones (1994) modified the Baek & Brock (1992) model, which relaxes the hypothesis that the series are iid, allowing them to show weak temporal dependence. Diks and Panchenko (2005) present a non-parametric test for Granger non-causality that avoids spurious rejection of the null hypothesis.

In the design of a bivariate test, the null hypothesis that past observations of $\{X_t\}$ do not contain additional information on Y_{t+1} is reformulated by Diks and Panchenko (2006) in terms of marginal distributions, as follows:

$$H_0 : \frac{f_{X,Y,Z}(x,y,z)}{f_Y(y)} = \frac{f_{X,Y}(x,y)}{f_Y(y)} \cdot \frac{f_{Y,Z}(y,z)}{f_Y(y)} \quad (4)$$

For simplicity, the authors consider $Z_t = Y_{t+1}$ and remove the time indices *t*. Thus, for each fixed value of y, X and Z are conditionally independent in $Y = y$. Thus, they show that another way of writing the null hypothesis is:

$$q \equiv E\left[f_{X,Y,Z}(X,Y,Z)\, f_Y(Y) - f_{X,Y}(X,Y)\, f_{Y,Z}(Y,Z)\right] = 0 \quad (5)$$

Next, taking $\hat{f}_W(W_i)$ as a local density estimate of a random vector d_W - variate $\boldsymbol{W}$ in $\boldsymbol{W_i}$, defined by $\hat{f}_W(W_i) = \frac{(2\varepsilon_n)^{-d_W}}{n-1} \sum_{jj\neq i} I_{ij}^W$, where $I_{ij}^W = I(\|W_i - W_j\| < \varepsilon_n)$, $I(\cdot)$ is the indicator function (that takes a value of 1 when true and null when false) and ε_n is the bandwidth parameter dependent on n (number of observations in the series); the Diks and Panchenko (2006) statistic test is:

$$T_n(\varepsilon_n) = \frac{n-1}{n(n-2)} \sum_i \left(\hat{f}_{X,Z,Y}(X_i, Z_i, Y_i) \hat{f}_Y(Y_i) - \hat{f}_{X,Y}(X_i, Y_i) \hat{f}_{Y,Z}(Y_i, Z_i)\right) \quad (6)$$

In the case of a bivariate test, the test is consistent; its power of rejection of a null hypothesis that is actually false tends to 1 (100%) as the amount of available data increases when $\varepsilon_n = Cn^{-\beta}$, where C is a positive constant and $\beta \in (1/4, 1/3)$. The asymptotic distribution to the normal pattern, in turn, is guaranteed in the absence of dependence between the $\boldsymbol{W_i}$ **vectors.** However, in a time series context, the authors demonstrate that asymptotic normality is guaranteed for weakly dependent data as long as the covariance between local density estimators is taken into account. The proof for this asymptotic property is developed under the "mixing" conditions of Denker and Keller (1983). In practice, we assume that taking $\varepsilon_n = Cn^{-\beta}$, with C being a positive constant and $\beta \in (1/4, 1/3)$, we have:

$$\sqrt{n}\frac{(T_n(\varepsilon_n) - q)}{S_n} \xrightarrow{D} N(0,1) \quad (7)$$

where $\xrightarrow{D}$ denotes convergence in the distribution and S_n is the estimator of the asymptotic variance of T_n.

Regarding the choice of ε_n, the *bandwidth* is optimal, in the sense of providing an estimator T_n with the smallest *Mean Squared Error* (MSE), when

$$\varepsilon_n = Cn^{-2/7} \qquad (8)$$

For practical applications, however, Diks and Panchenko (2006) recommend a truncation, as follows:

$$\varepsilon_n = min\left(Cn^{-2/7};\ 1.5\right) \ (9)$$

That is, the value of ε_n is truncated to hold the maximum at 1.5 to avoid obtaining very large bandwidth values in the case of a small number of available observations (n) for the series. In our study, we choose a conservative value and use $\varepsilon_n = 1$, in line with other studies (Andreasson et al., 2016; S. D. Bekiros & Diks, 2008b).

2.5
Results

Considering the approaches described, we present the results of the linear and nonlinear tests for each of the variables, the T index of speculation, and the long and short positions of the CITs below. The following tables show the relationships investigated.

Table 2 presents the results of the Toda and Yamamoto (1995) method before and after the filtering to control conditional heteroscedasticity.

Even with the versatility of the models, heteroscedasticity can affect the series. Thus, we use the equations using ARCH family filters. We find a characteristic behavior in the financial market when periods of high volatility are preceded by low volatility, and vice versa. Thus, the ARCH family models can better capture this stylized fact; the variance of the present error term is related to the size of the error of the previous period. We can test the conditional distribution of the error term using the normal distribution (Gaussian), Student's *t*, and the generalized error distribution. The choice of the best model is given by the statistical significance of the parameters and the values of information criteria, such as AIC and Bayesian Information Criterion.

The results in Table 2 indicate that there is no causal relationship between the funds and speculation towards the price in most commodities. However, the CIT sell position for sugar and wheat point to causality, in the direction of price; that is, the sell position may have influenced prices.

	Toda Yamamoto Causality			GARCH filtered residuals		
Commodities	CIT_L	CIT_S	T	CIT_L	CIT_S	T
Coccoa→ Price	-	-	-	-	-	-
Price→ Coccoa	-	*	-	***	-	-
Cotton→Price	-	-	-	-	-	-
Price→Cotton	-	-	-	**	*	-
Feeder Cattle→ Price	-	-	-	-	***	-
Price→ Feeder Cattle	-	-	***	-	***	***
Lean Hogs→Price	**	-	-	-	-	-
Price→ Lean Hogs	-	-	***	-	-	-
Live Cattle→Price	-	-	-	-	**	-
Price→ Live Cattle	***	-	***	***	-	-
Coffee→ Price	-	-	-	-	-	-
Price→ Coffee	-	-	***	-	***	-
Corn→ Price	-	-	-	*	-	-
Price→Corn	*	*	***	*	-	-
Soybean→ Price	-	-	-	-	-	-
Price→ Soybean	***	-	*	***	***	-
Sugar→Price	*	**	-	-	**	-
Price→ Sugar	-	-	***	-	-	-
Wheat→ Price	-	***	-	-	***	-
Price→ Wheat	-	-	***	-	-	-

Table 2. TY Causality

Note: Toda-Yamamoto procedure (Modified Wald statistics). Commodity (X) → Price (Y) denotes the position of CITs, and Granger speculation (T) does not cause price, or vice versa. *, **, and *** show significance at the 10%, 5%, and 1% level, respectively. The "*GARCH filtered residuals*" column shows the results of the TY procedure using the residuals filtered by appropriate GARCH models. The size of lags was determined by the BIC and AIC criteria, with a maximum of 6 lags.

The choice of the *lead-lag* plays an important role in the application of causality tests. In Tables 2 and 3 (Standard Granger), to determine the optimal number of lags of the bivariate VAR, the maximum number of lags is set to 6 and the optimal number is selected by means of some AIC or BIC (in this case, we choose it by the smallest number of lags for parsimony).

	Granger-Wald (VAR)			BEKK-GARCH Filtered Residuals		
Commodities	CIT_L	CIT_S	T	CIT_L	CIT_S	T
Cocoa→ Price	-	-	-	-	-	-
Price→ Cocoa	-	*	-	**	*	***
Cotton→Price	-	-	-	-	-	-
Price→Cotton	**	-	-	**	-	-
Feeder Cattle→ Price	-	-	-	*	*	-
Price→ Feeder Cattle	-	-	***	-	-	***
Lean Hogs→Price	**	-	-	***	-	**
Price→ Lean Hogs	-	-	***	*	**	***
Live Cattle→Price	-	-	-	-	***	*
Price→ Live Cattle	**	-	***	***	-	***
Coffee→ Price	-	-	-	-	-	-
Price→ Coffee	-	-	***	-	-	***
Corn→ Price	-	-	-	*	-	-
Price→Corn	-	*	***	-	-	-
Soybean→ Price	-	-	-	-	-	**
Price→ Soybean	***	-	*	**	-	**
Sugar→Price	-	*	-	-	**	**
Price→ Sugar	-	-	***	-	-	-
Wheat→ Price	-	***	-	-	***	-
Price→ Wheat	-	-	***	-	-	***

Table 3. Granger standard causality (Wald-Granger)

Note: Granger causality test. Commodity (X) → Price (Y) denotes position of CITs, and Granger speculation (T) does not cause price, or vice versa. *, **, and *** show significance at the 10%, 5%, and 1% level, respectively. The column BEKK Filter shows the residuals of VAR filtered using the BEKK method. Size of lag was determined by the BIC and AIC levels, with a maximum of 6 lags.

Table 3 shows the results of the traditional GC test using a bivariate VAR for the log-returns of the involved variables. Next, we filter the residuals using the BEKK-CARCH multivariate model. A formulation that can be viewed as a restricted version of the VECH model is a specific parameterization of the multivariate GARCH model defined in Engle and Kroner (1995). This formulation is called BEKK, an acronym that stands for Baba, Engle, Kraft and Kroner (Baba et al., 1990; Engle & Kroner, 1995), in line with recent studies on nonlinear causality (Andreasson et al., 2016; S. D. Bekiros & Diks, 2008a), and has the following form:

$$H_t = C'C + \sum_{i=1}^{q} A_i' \epsilon_{t-i} \epsilon_{t-i}' A_i + \sum_{j=1}^{p} B_j' H_{t-j} B_j, \qquad (13)$$

where now C, A_i and B_j are N x N matrices, C being lower triangular. In this model the conditional covariance matrices are positive definite by construction, which is an attractive property.

Before and after filtering using BEKK-GARCH, Table 3 indicates that there is no relationship between speculation and the commodity prices. That is, speculation did not cause price variation for most commodities. However, there is one relationship, in the Granger sense, from speculation to price. Thus, our findings suggest that prices guide speculators (*trend followers*). This is in line some prior studies (Guilleminot et al., 2014; Lehecka, 2015). Furthermore, as Table 2 shows, the sold positions of the CITs seem to have caused price variation for sugar; with filtering, the significance level increases.

	DP (2006)			BEKK-GARCH Filtered Residuals		
Commodities	**CIT_L**	**CIT_S**	**T**	**CIT_L**	**CIT_S**	**T**
Coccoa→ Price	-	-	-	-	-	-
Price→ Coccoa	-	-	-	-	-	-
Cotton→Price	-	-	-	-	-	-
Price→Cotton	-	-	*	-	-	-
Feeder Cattle→ Price	***	-	**	***	-	*
Price→ Feeder Cattle	-	-	***	-	-	**
Lean Hogs→Price	-	-	*	-	-	-
Price→ Lean Hogs	-	-	**	-	-	-
Live Cattle→Price	**	-	-	**	-	-
Price→ Live Cattle	*	-	**	-	-	-
Coffee→ Price	-	-	-	-	-	-
Price→ Coffee	-	-	**	-	-	-
Corn→ Price	-	-	-	-	-	-
Price→Corn	-	-	***	-	-	-
Soybean→ Price	-	-	-	-	-	-
Price→ Soybean	-	-	-	-	-	-
Sugar→Price	**	-	-	**	-	-
Price→ Sugar	-	-	-	-	-	-
Wheat→ Price	-	*	-	-	-	-
Price→ Wheat	-	-	**	-	-	-

Table 4. Diks and Panchenko nonlinear causality (2006)
Note: Nonlinear causality DP (2006) T_n statistic. Commodity (X) → Price (Y) denotes position of CITs, and Granger speculation (T) does not cause price, or vice versa. *, **, and *** show significance at the 10%, 5%, and 1% level, respectively. Bandwidth set at $\boldsymbol{\varepsilon_n = 1}$ and lag ℓ=1. VAR residuals filtered using a BEKK-GARCH multivariate specification.

With the data filtered using the multivariate BEKK-GARCH model, the use of the Diks and Panchenko (2006) nonlinear test makes it possible to determine whether the model can describe the relationship among the analyzed commodities. Filtering is applied when there is a failure to accept the null hypotheses of non-causality of the nonlinear model (S. Bekiros, 2014).

The test consists of applying the DP model to the stationary series to detect whether there is some nonlinear relationship. Following this step, the filtered residuals of the VAR-BEKK model are generated to verify whether the nonlinear relationship persists. We set both bandwidth and lag at 1. According to S. D. Bekiros & Diks (2008a), if the bandwidth is defined as having a value greater than (or less than) 1, it will result in a smaller (or larger) p-value.

As Table 4 shows, most of the results were inconclusive, and after filtering, the significance of the relationships diminished. In contrast to these findings, we note that a buy position on the part of the CITs generated pressure on the price of sugar, feeder cattle, and live cattle, both before and after the filtering. This confirms that statistical significance is weaker after filtering for all commodities.

Considering our results, the hypothesis that speculators and the *index funds* caused price hikes may not be accepted considering our tests. However, one must exercise caution in admitting the hypotheses that the agents are *trend followers* given that the nonlinear model is more precise in capturing relationships over time, and the results show no such relationship after filtering.

2.6
Conclusion

Using different approaches, we illustrate whether there is any relationship between speculation and the position of the funds in the prices of ten commodities extracted from the CFTC report. We tested the hypothesis that speculators and *index funds* were responsible for the pressure on agricultural commodity prices, both by models that capture the linear relationship and by those seeking to indicate possible nonlinear relationships.

The results suggest that speculation did not cause changes in agricultural commodity prices. Linear models point precisely in the opposite direction, i.e., that prices Granger cause speculation; that is, speculators seem to be guided by price changes. They buy when the price increases and sell when the price falls (*trend followers*). In fact, the nonlinear test was inconclusive, with no consistent relationship in either direction.

In general, the results indicate that causal relations in both directions are weak. However, caution should be exercised in testing since the data disclosed may distort relationships. Among these distortions is the fact that the position is disclosed only weekly; that is, within weeks, there may be important movements that the analysis does not include. In addition, large investment funds may use different strategies in different markets and may also operate *intraday* or take some position without apparent cause and effect, which masks the pattern of causation. Large traders who operate as *hedgers* may take a speculative position. (Irwin et al., 2016) point out that *hedgers* act as speculators with exclusive information in some cases. Another factor to consider is that the period prior to 2006 is not disclosed in the SCOT report—which may limit the test.

Therefore, with the data available at the time of this publication, it is not possible to prove the effect of speculation on commodity prices, even with robust methodologies. Although the findings point to no causality, one may not conclude that speculation did not play a role in price pressure, or vice versa. There is still a gap in the dissemination of data in these markets.

This study highlights findings that are relevant to financial market regulators, *traders*, and commodity producers. We suggest that future studies should aim to address the question of whether there was a herd or cascade behavior that caused prices to escalate, and whether the trigger was any reason besides speculation.

3
The sugar-ethanol-oil nexus in Brazil: Exploring the pass-through of international commodity prices to national fuel prices

3.1
Introduction

Fuel, particularly ethanol and gasoline, prices constitute an important driver of the Brazilian economy. Concerning ethanol, not only the country is a leading global producer and net exporter, but it also possesses the world's largest fleet of ethanol-fuelled cars. The US and Brazil together supply 85% of the ethanol worldwide (see Figure 2. In 2019, Brazil undertook the second-largest expansion in ethanol production in the world. According to the Organization for Economic Co-operation and Development and Food and Agriculture Organization (OECD, Food, & of the United Nations, 2019a), ethanol use is expected to rise roughly 18% by 2028, and Brazil stands as one of the countries that could supply this additional output. Gasoline prices, in turn, have a significant weight on inflation in Brazil as they account for a considerable share of transportation costs. In addition, gasoline prices also exert considerable influence on ethanol commercialization, given the parity, i.e., the price ratio equilibrium in the gasoline–alcohol blend, and the widespread dissemination of the flex-fuel technology in internal combustion engines. In this context, understanding the manner and the extent of the transmission mechanisms from international energy and commodity prices to Brazilian fuel prices is crucial for effective policy making and optimal risk management strategies. Furthermore, not having efficient risk-sharing mechanisms during periods of considerable price volatility may distort input allocation, curb agricultural investment and slow productivity growth. This could have severe implications for consumers, farmers and countries that depend on biofuels (Ceballos, Hernandez, Minot, & Robles, 2017).

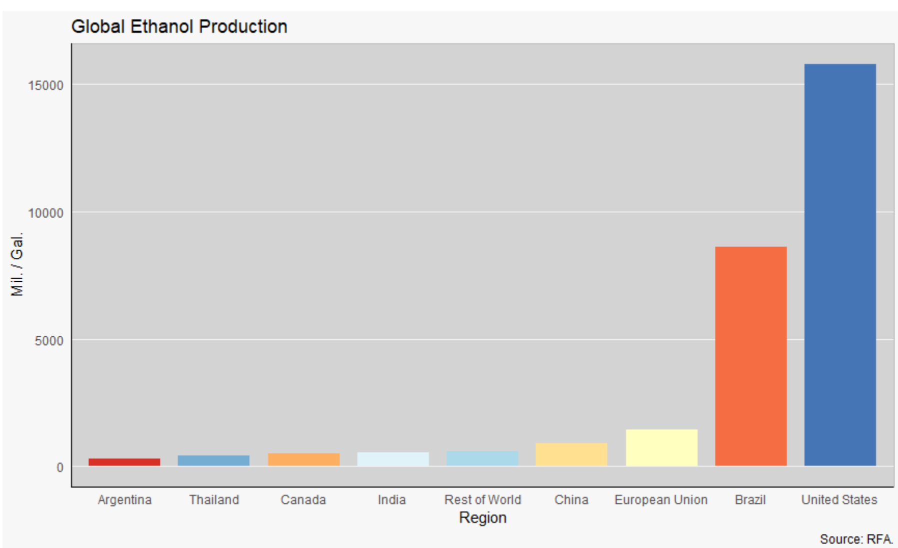


Figure 2. World ethanol production in 2019 (million gallons). Source: RFA (2020).

Understanding fuel price dynamics in Brazil, however, is not straightforward, particularly given that commodities markets have undergone a profound transformation since the early 2000s. Throughout this process, named the financialisation of commodity markets, there have been huge spikes in prices simultaneously across different markets (e.g. energy, agriculture and metals), and a great deal of investments has been channelled to these markets (Cheng & Xiong, 2014; Irwin & Sanders, 2011). A stream of the literature highlighted the abrupt change in price behaviour and the main factors behind the booms and slumps in commodity markets. For example, Figuerola-Ferretti et al. (2015) showed evidence of mild explosivity in the metals prices in 2004-2007. They explored whether price movements were related to fundamentals or speculation. In the same strain, Figuerola-Ferretti et al. (2020) analysed the potential drivers of the 2007-08 oil price spike and the 2014-16 price drop. They detected a mildly explosive episode during 2007-08, suggesting that the rise in oil prices was not driven by speculation but by global demand. The food price boom also was examined. Gilbert et al. (2020) attributed the 2007-08 price spike in grains to the limited refining capacity of the US to meet the demand for ethanol created by renewable fuel mandates. Mitchell (2008) ascribed the rise in trade food prices to the expansion of biofuels production.

The Brazilian fuel industry also experienced market transformations. The Brazilian government used administered prices to control inflation, mainly between 2011 and 2014, and this exerted a direct influence on gasoline prices (Almeida et al., 2015). However, in 2015, the state-controlled oil company, Petrobras, was forced to increase fuel prices to reduce its debt burden, and in 2016 the former president of Petrobras declared that gasoline prices would follow the global trend in oil prices. Thus, we posit that the prices of international commodities became more likely to influence and pass through to domestic prices. Government pricing policies can delay the convergence of global prices to domestic prices. For example, in 2007 and 2008, many countries raised their subsidies by not passing through the entire rise in international oil prices. The absence of mechanisms to transmit international prices to domestic markets may affect global demand. Lower domestic prices increase demand, and this can promote an increase in the international price (Cavalcanti, Szklo, & Machado, 2012; Piotrowski et al., 2010).

Despite Brazil's strategic role in the international ethanol market, there are no international proxies for Brazilian ethanol prices. In addition, the known factors that drive Brazilian ethanol prices are mostly domestic (David, Inácio, Quintino, & Machado, 2020). The first is the idiosyncratic nature of sugarcane activity, such as the price inelasticity of demand. The second concerns the rigidity of short-term supply related to the intertemporal decision about production, which is intrinsic to price uncertainty. Third, gasoline prices, the supply of sugarcane, the balance of production, and the number of flexible-fuel cars are domestic factors related to ethanol price volatility (David, Quintino, Inacio, & Machado, 2018).

In the context of the above, several questions remain. Do international energy and feedstock prices exert influences on Brazilian ethanol and gasoline prices? If so, which markets influence them the most, and what is the extent of these influences? Are they symmetric or asymmetric according to the direction (negative or positive) of oil price shocks? Do they differ according to the horizon (short- or long-term) considered? The motivation lies in the long discussion concerning the asymmetrical effects of crude oil prices on gasoline market price behaviour, named "rockets and feathers" by Bacon (1991), who argued that oil prices would rise like a rocket but fall like a feather.

To address such questions, the use of nonlinear models to investigate price dynamics is timely, given the intricacies of the commodity markets, such as sudden policy shifts, regime-switching behavior, asymmetric responses to news, financial turbulence, and leverage effects (Erick Meira de Oliveira, Cyrino Oliveira, Klötzle, & Pinto, 2019). As a result, gasoline and ethanol prices tend to have a nonlinear relationship with energy commodity prices (Hammoudeh, Lahiani, Nguyen, & Sousa, 2015). Recent authors have argued that nonlinear approaches provide a better understanding of the dynamics of commodity prices. For instance, Kwek and Koay (2006) argued that symmetric models are not adequate to explain exchange rate behaviour. Nazlioglu (2011) compared the linear and nonlinear relationships between oil and agricultural commodities, and he acknowledged that the nonlinear approach provided a better comprehension of the commodities' price dynamics. Atil et al. (2014) examined asymmetric price transmission between energy commodities. They showed that the behaviour and interaction of economic and financial variables over time are mostly driven by nonlinear patterns. In this paper, multivariate nonlinear autoregressive distributed lag (NARDL) models are built using weekly prices of energy commodities traded in futures markets: heating oil (HO), reformulated blendstock for oxygenate blending (RBOB), West Texas Intermediate (WTI) crude oil, Chicago ethanol (Platts), corn and sugar—as independent variables—and Brazilian ethanol and gasoline prices as the dependent variables. The period of analysis runs from January 2015 to May 2020.

Our main findings show that: (i) RBOB has an asymmetric effect on Brazilian gasoline prices, (ii) surges in HO lead to a decrease in ethanol prices in the long-term, and (iii) there is unidirectional nonlinear causality from HO to ethanol prices. The novelty of this work lies not only in our findings but also in how we explore the interplay between energy and agricultural commodities. Previous studies have overlooked the interaction between international energy and feedstock prices and domestic fuel prices in a multivariate framework. In addition, few studies have focused on the pass-through mechanisms of global commodity prices to ethanol or gasoline prices in Brazil.

The rest of the paper is organised as follows. Section 2 provides an overview of the Brazilian fuel market and its development, and reviews the related literature. Section 3 describes the data and methodology. Section 4 presents and discusses the empirical results. Section 5 concludes and presents suggestions for future studies.

3.2 Background

3.2.1 The Brazilian fuel Market

This section presents the salient features of the Brazilian domestic fuel industry, the main characteristics of fuel products and how ethanol and gasoline prices are set in the domestic context. Acknowledging the basic domestic price mechanisms is key to further understand the oil–sugar–ethanol nexus in Brazil.

In 2019, Brazil accounted for the second-largest expansion in ethanol production (OECD, Food, & of the United Nations, 2019b). In the 2018–2019 crop season, ethanol achieved its highest dominance in the sugar production mix. It should be noted that most mills in Brazil have the flexibility to produce either sugar or ethanol, depending on market conditions (S. M. de Oliveira, Ribeiro, & Cicogna, 2018). This flexibility of producing sugar and ethanol (mix) revolves around 40 to 60% (McKay et al., 2016). Thus, the price dynamic in the international sugar market is an important driver to switch between sugar and ethanol output.

In Brazil, ethanol is produced mainly from sugarcane, which accounts for 99% of the national ethanol production. Over 90% of the total sugarcane harvest takes place in Brazil's South-Central region. Sugarcane ethanol production in Brazil has a relative advantage, because of its efficient use of land, compared to corn ethanol production in the US. However, transportation costs in Brazil, along with the value of currency and pricing of co-products, can alter this competitive advantage (Crago et al., 2010).

Brazil is the largest sugar exporter and the second ethanol exporter in the world (OECD et al., 2019a). Brazil had been the world's leading ethanol producer until the US surpassed its production in 2005 (Araújo, 2016). Nevertheless, Brazil is still a relevant consumer of ethanol. More than 80% of its fleet vehicles are flexible-fuel (ANFAVEA, 2020). Not only can Brazilian consumers choose between gasoline and ethanol, but they can also change the proportions of these products in the same tank. Over the last five years, this convenience has led to an increase in the demand for hydrous ethanol from owners of flexible-fuel cars.

There are two types of ethanol: anhydrous, which is used to blend into gasoline, and hydrous, used directly in vehicles (E100). Different tax regimes for ethanol usually favour the production of the latter (OECD et al., 2019b). Currently, the national mandate, which changes by government decree, requires 27% ethanol in the gasoline–alcohol blend (gasohol). It is worth noting that from early 2015, this proportion increased from 25% to 27%. The choice between the two types of fuel is usually based on the price ratio—the parity. The parity equilibrium is equivalent to 70% ethanol/gasoline (Furtado, Scandiffio, & Cortez, 2011). This means that above this ratio, E100 is no longer competitive. Furthermore, the gasoline price ceiling dictates the price of hydrous ethanol, while the supply from sugarcane mills and the distributors' demand for blending gasoline determine the anhydrous price (S. M. de Oliveira et al., 2018). Moreover, Melo and Sampaio (2014) showed that ethanol consumers' demand increased because of the rise in the gasoline price, so consumers tend to switch to ethanol rather than gasoline in the long run.

Petrobras, a national state-owned company in Brazil, explores, produces, refines, and distributes oil products to wholesalers through its retail company, Petrobras Distribuidora S.A. Figure 3 shows the structure of the Brazilian fuel industry. Gasoline A (ethanol-free) can be produced either by Petrobras or other producers (refineries and formulators) or imported. Distributors (C) buy gasoline A from Petrobras' refineries (A) and ethanol from sugar mills (B). Distributors blend gasoline A with anhydrous ethanol to formulate grade-C gasoline to sell to gas stations. The refineries establish the gasoline price, and their costs are determined by the inputs and the structure of the refinery plus the net margin of the refiner or importer and taxes. Then, wholesalers buy gasoline A and sell gasoline C to retailers (gas stations), adding their margin taxes/subsidies[1] (Cavalcanti et al., 2012). In our framework, we consider gasoline type A and hydrous ethanol as the dependent variables.

Petrobras is one of the largest oil producers in the world, but it cannot refine enough to meet all the local demands. Thus, Petrobras and other agents must import oil products such as diesel and type A gasoline. Notably, the country logistics depend mostly on road transport, implying a high diesel consumption (Derick David Quintino & Ferreira, 2021). On top of that, diesel is the highest imported oil product in Brazil, and approximately 85% of imported diesel comes from the US. Until 2014, Petrobras was the main importer of diesel and gasoline. This situation changed in early 2015 when private companies were allowed to import diesel and gasoline. As a result, the importation shares of the private agents jumped from 0% in 2014 to 16% in 2015 and 72% in 2016. Yet, only three companies hold sway over more than 50% of the distribution market: Petrobras, Raízen and Grupo Ultra.

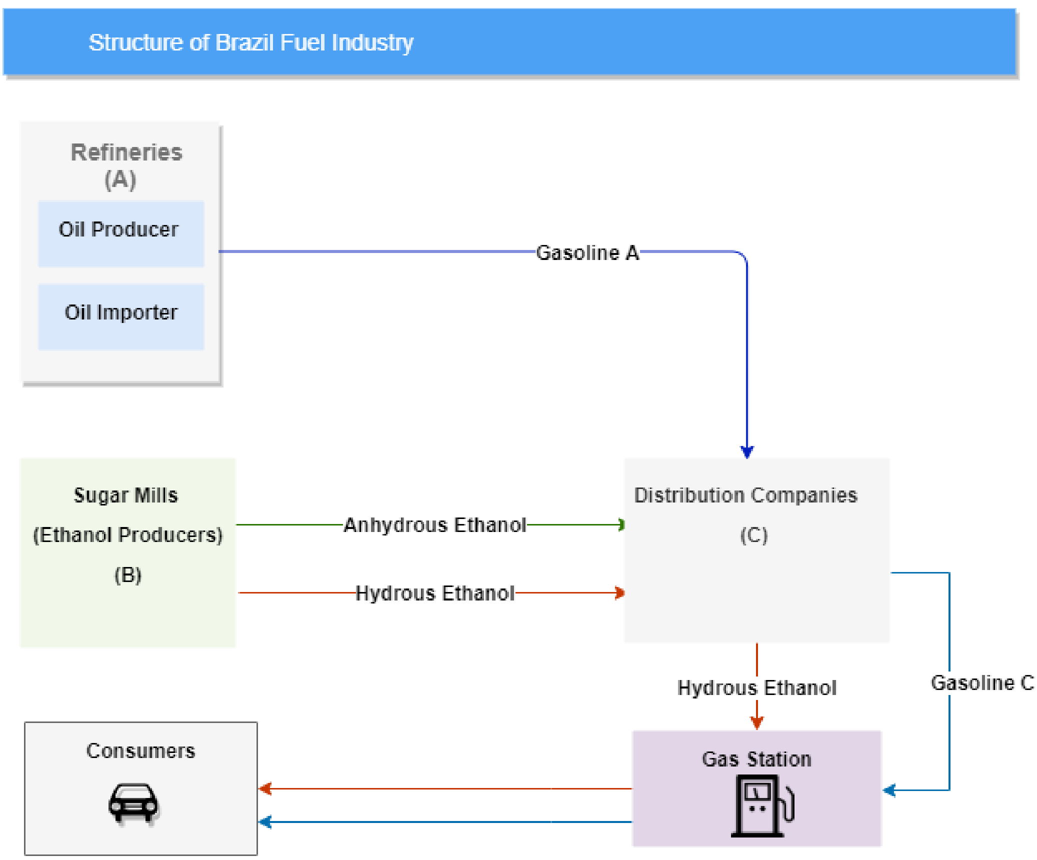


Figure 3. The structure of the Brazilian fuel industry and the distribution process.

3.2.2
The structural changes in the Brazilian fuel market

[1] See https://petrobras.com.br/en/our-activities/composition-of-sales-prices-to-the-consumer/gasoline/for the overall consumer price composition.

To understand the structural changes in the fuel sector in Brazil, it is worth noting the political context before 2015. In the first term of President Dilma, from 2011 to 2014, the government adjusted prices at its discretion to meet the inflation target. At the beginning of the second term in 2015, however, that policy proved to be ineffective, leading to price distortions (de Azevedo & Serigati, 2015). Economic measures in the first term, such as fixed prices for gasoline and electricity, did not linger in the second term. Furthermore, the finance minister launched a series of austerity measures to stabilise and reduce the gross public debt, raising electricity, fuel, and other regulated prices[2]. Thus, despite the decline in international crude oil prices after 2015, the increase in administered prices and the depreciation of the Brazilian currency created additional inflationary pressure. Even though global oil prices decreased considerably, Petrobras was forced to increase its fuel prices mainly to reduce its debt burden[3]. This contributed to the increase in hydrous ethanol commercialisation.

According to Nascimento Filho et al. (2021), price control was possible because of two reasons. First, because Petrobras has been the only fuel supplier for Brazilians, being responsible for more than 80% of the market share for Gasoline A delivered to the distributor chain by 2019 (ANP, 2019). Second, given the fact the Brazilian federal government holds most of the ordinary shares of Petrobras (PETR3). Nascimento Filho et al. (2021) add that there was also a lack of definition regarding who was mainly responsible for the Petrobras' pricing policy. As a result, the variation in international prices was applied with a delay to fuel prices in the country.

The situation changed considerably in 2015 when Petrobras promoted a new diesel and gasoline company pricing policy, based on Import Parity Price (IPP) plus a price margin to remunerate the risks inherent to the operation. The new pricing policy increased the importance of markets when setting fuel prices (Hallack et al., 2020), which contributed to considerably higher levels of volatility in gasoline prices. Significant changes were also observed in ethanol–gasoline price ratio behavior after the promotion of the IPP, with nonlinear behavior, reported on several occasions (Nascimento Filho et al., 2021).

Even though Petrobras started to adjust prices in 2015, it was only in June 2016 that the recently appointed President of Petrobras announced that gasoline prices would be guided by the company's own interests, without government interference. Petrobras thus began adjusting the fuel prices regularly, following international trends in oil prices. However, the regular adjustments could not track the high volatility in oil prices and the Brazilian currency rate. Thus, in June 2017, price adjustments occurred more frequently, sometimes even daily. The policy of adhering to global movements in oil prices lasted until May 2018, when the trucker strike hit the Brazilian economy. The lorry-driver strike caused turmoil in the Brazilian economy, resulting in shortages of food and fuel. Then, the government succumbed to the lorry driver's pressure and reduced diesel prices. In the following months, Petrobras re-established price adjustments according to international parity, but with a lower frequency.

[2] See https://foreignpolicy.com/2015/06/01/brazils-economy-dilma-rousseff-makes-amends-with-the-markets/.
[3] The fuel subisidies had cost more than $24bn to Petrobras. See https://www.ft.com/content/ea974c28-669e-11e4-91ab-00144feabdc0.

One side effect of the government-mandated fuel price ceilings was the impact on the sugarcane sector, specifically on ethanol production and investment. Costa et al. (2016) showed that gasoline price control significantly harmed hydrous ethanol production. As parity increased above the equilibrium (approximately 70%), fewer consumers were willing to pay for ethanol, culminating in a sharp fall in demand (Salles-Filho et al., 2017). A caveat is that gasoline prices were constantly held above international prices from 2006 to 2010, benefiting E100 production (Costa & Burnquist, 2016).

To compensate for the industry downturn, the Brazilian government altered the ethanol mix in gasoline from 25% to 27%, and a government program called RenovaBio stimulated ethanol production. Hydrous ethanol sales increased by 42% from 2018 to 2019. The Brazilian Ministry of Mines and Energy implemented the RenovaBio program to reduce greenhouse gas emissions under the Paris Agreement and dependency on diesel and gasoline importation and to increase the contribution of biofuels in the Brazilian energy matrix (ANP, 2019; Branco et al., 2019). The RenovaBio program established a 10-year target to reduce greenhouse gas emissions; every ethanol producer would be certified by its total CO_2 emission in the production process. The Brazilian National Agency for Petroleum, Natural Gas and Biofuels (ANP) is responsible for determining how many litres of biofuel are needed to prevent the emission of one ton of CO_2 into the atmosphere. This amount is equivalent to a Biofuel Decarbonisation Credit (CBio). Thus, CBio would be used as a financial instrument to trade carbon credits in the market to offset the distributors' CO_2 emissions.

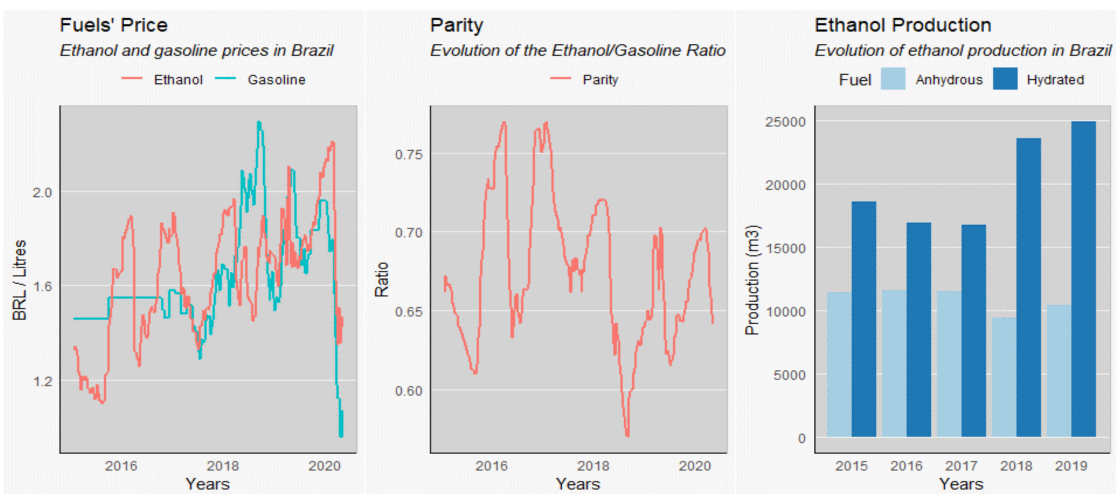


Figure 4. Ethanol (BRL/1,000 l) and gasoline prices (BRL/l), parity (ethanol/gasoline ratio) and ethanol production (m3). Sources: (ANP, Brazilian National Petroleum, 2020; Cepea, 2020).

3.2.3
Literature Review on Oil–Gasoline Price Dynamics

The asymmetrical effects of crude oil prices on the market price of gasoline have long been a subject of discussion among scholars. Bacon (1991) named this asymmetrical pattern 'rockets and feathers', meaning that oil prices would rise like a rocket but fall like a feather. The author provided evidence of the 'rockets and feathers' pattern for the speed of adjusting UK retail gasoline prices to cost. Other studies also have supported the rockets and feathers idea. For example, Borenstein et al. (1997) supported the common belief that crude oil prices had a faster effect on increases than on decreases in retail gasoline prices in the US. Galeotti et al. (2003) found that gasoline price adjustments to oil increases were faster than oil price reductions in five different European countries. Polemis and Tsionas (2016), using a nonlinear semi-parametric model, found evidence of asymmetric oil price adjustments for wholesale and retail gasoline prices. Moreover, using a NARDL approach, Atil et al. (2014) and Kisswani (2019) corroborated the asymmetric effects of oil prices on gasoline prices.

The asymmetric responses of gasoline prices to oil price movements have different explanations. For instance, Borenstein et al. (1997) argued that market power and oligopolistic coordination with imperfect monitoring explained the price adjustments. Radchenko (2005) pointed out that oligopolistic coordination was the primary cause of the asymmetry. Findings from other studies suggested a positive link between market power and asymmetric pass-through (Farkas & Yontcheva, 2019; Oladunjoye, 2008). Conversely, Borenstein et al. (2002) argued that refiners with market power adjusted prices slowly in either direction—when oil prices increased or decreased. Bumpass et al. (2015) showed that retail and wholesale gasoline prices change symmetrically with oil price shocks in the long run. Honarvar (2009) attributed the main effect on gasoline prices to technological changes in demand instead of oil price movements per se, as suggested by the literature.

3.2.4
International Dynamic Relationship between Oil, Biofuels and Agricultural Commodities

Knowing whether international commodity prices co-move is key to understanding the relationship between oil, biofuels and agricultural commodities. Pindyck and Rotemberg (1990) argued that commodity prices tend to move together, regardless of macroeconomic variables; they attributed co-movements to herd behavior[4]. Macroeconomic variables, such as aggregate demand, inflation and exchange and interest rates, may generate co-movements in nominal commodity prices. In addition, microeconomic factors, such as substitutability or complementarity, may explain the correlation between certain commodities. However, complementary commodities or substitutes may exhibit a negative price correlation (Leybourne et al., 1994). Nevertheless, since the early 2000s, several investment funds have emerged in the commodity markets, seeking a new asset class. The so-called financialisation process in commodity markets has forced non-energy commodity prices to become more correlated with oil prices. Consequently, commodity prices are not determined by supply and demand only but also by the risk-return appetite of investors and their portfolio diversification (Tang & Xiong, 2012).

After the 2007–2008 boom in commodity prices, several studies tried to explain the reasons for the above phenomena. The main factor for price changes was debated; factors such as exchange rates and supply and demand, as well as speculation, were analyzed (Abbott, Hurt, & Tyner, 2008; Palazzi et al., 2020; Sandersrs & Irwin, 2011; Timmer, 2008; Trostle, 2008). Mitchell (2008) blamed the increase in biofuel production for the spike in internationally traded food prices. Wright (2014; 2011) claimed that biofuel policies led to a shift in grain prices. Gorter et al. (2016) argued that developing countries' policy responses had little impact on world food prices in 2008 and that the 2007–2008 boom induced a structural shift in the market that everybody overlooked.

Moreover, the commodities market dynamic drove the demand for biofuel production to be more related to agricultural markets as oil prices increased—a surge in agricultural commodity prices was attributed to changes in oil prices. Because of the increased use of soybean and corn to produce ethanol and biofuel, oil price linkages became tighter (Nazlioglu, 2011). In this sense, Nazlioglu et al. (2013) showed that during the pre-crisis period (1986–2005), there were no risk transmissions between oil and agricultural markets. By contrast, the variance causality test showed a spillover from oil to agricultural markets in the post-crisis period. In Pakistan, where petroleum is predominantly imported, Sarwar et al. (2020), using NARDL showed that oil prices have a positive asymmetric effect on food and non-food inflation.

3.2.5 The Sugar–Ethanol–Oil Nexus

As previously outlined, Brazil and the US are not only the primary producers of ethanol, but they are also significant consumers. The demand for biofuels has expanded considerably, driven mainly by rising concerns about environmental issues. In Brazil, particularly, production has been spurred by a variety of measures, such as government policies (e.g., RenovaBio), technological changes and the growth of the flexible-fuel fleet.

[4]See Júnior et al. (2019) for a detailed discussion about herding behavior in commodities markets.

In this connection, it is safe to assume that ethanol price behaviour in Brazil has become an important driver of the national fuel industry dynamics, and, therefore, the effects of the former on the latter should be closely investigated.

The sugar–ethanol–oil nexus also helps to explain the complexity of the Brazilian fuel market. Thus, the hypothesis of oil driving gasoline prices is not straightforward when ethanol is included. In this context, Balcombe et al. (2008) found that oil prices determine the long-term equilibria of sugar and ethanol prices in Brazil, and the leading order was oil to sugar to ethanol. According to the framework proposed by Serra et al. (2010), crude oil and sugar are exogenous factors—an increase in oil prices drives an upward change in ethanol prices, and an increase in sugar prices increases ethanol prices' levels and volatility. On the other hand, Brazilian policies can affect ethanol and corn prices in the US (Drabik, De Gorter, Just, & Timilsina, 2014), and biofuels can also affect the prices of agricultural products (Maitah, Procházka, Smutka, Maitah, & Honig, 2019).

As outlined, many studies have explored the relationship between oil and gasoline prices in the international markets and the price response of agricultural commodities to shocks in oil prices. Nevertheless, the results have so far not been conclusive about the asymmetric effects on the dependent variables. Moreover, few studies have analysed the Brazilian ethanol and gasoline markets, which are essential players in global commodities trading. There is also a lack of studies investigating the impacts of oil and feedstock prices on ethanol and gasoline prices in Brazil. The present study attempts to fill these gaps by assessing the mechanisms of asymmetric pass-through from international oil and agricultural prices to Brazilian ethanol and gasoline prices.

3.3
Data and Methodology

3.3.1
Data

We considered weekly closing prices between January 2015 and May 2020 of the main energy commodities traded in the international futures markets and the two major agricultural commodities used in ethanol production as independent variables. The criteria for selecting futures contracts are related to the importance of these contracts as international price benchmarks. The dependent variables were the Brazilian ethanol and gasoline spot prices. The price of gasoline (type A) refers to the tax-free price delivered by Paulínia–São Paulo, the main refinery in Brazil (described in section 2.2). Although there is a hydrous ethanol future contract traded on the Brazilian stock exchange (B3), the spot market is the dominant market between futures and spot prices. Thus, according to Quintino et al. (2017), the dominance in the cash market is attributed not only to low liquidity but also to the market's concentration in the ethanol wholesale market and the oligopolistic behaviour of the distributors. Particularly, Grupo Ultra is the second-largest distributor, followed by Raízen, which is not only a major distribution company but is also one of the largest ethanol producers in Brazil. For instance, both companies designated the use of HO and RBOB futures contracts (long/short position) to hedge their margin exposures reported in their financial statements. This suggests that HO and RBOB play important roles in price formation for refinery.

Table 5 illustrates the specifications of the variables involved, while Table 6 provides the descriptive statistics for their log-transformed versions. The nearby futures contracts (independent variables) were retrieved from Bloomberg. Each futures contract was rolled over on the last trading days.

Variable	Notation	Market	Price Quotation	Source	Ticker
West Texas Intermediate Crude Oil	WTI	CME Group	U$ cents/barrel	Bloomberg	CL1 Comdty
Heating Oil	HO	Nymex	U$ cents/gallon	Bloomberg	HO1 Comdty
Reformulated Blendstock for Oxygenate Blending	RBOB	CME Group	U$ cents/gallon	Bloomberg	XB1 Comdty
Chicago Ethanol (Platts)	EPlatts	CME Group	U$ cents/gallon	Bloomberg	CUA1 Comdty
Sugar #11	Sugar	The ICE	U$ cents/pound	Bloomberg	SB1 Comdty
Corn	Corn	CME Group	U$/bushel	Bloomberg	C 1 Comdty
Brazilian gasoline type A (Refinery) *	Gasoline	Spot	US$/liter	ANP	-
Brazilian Ethanol (Hydrous) prices*	Ethanol	Spot	US$/liter	ESALQ/CEPEA	-

Table 5. Selected dataset
Note: Weekly prices collected from January 2015 to May 2020. * Represent the dependent variables in US dollar. Conversions from BRL to USD were made using the official exchange rate PTAX from the Brazilian Central Bank.

	Corn	EPLATTS	Ethanol	Gasoline	HO	RBOB	Sugar	WTI
Log-levels								
Mean	5.92	0.36	6.11	-0.80	5.13	5.07	2.64	3.93
Median	5.92	0.38	6.12	-0.77	5.17	5.10	2.60	3.96
Maximum	6.14	0.59	6.40	-0.53	5.48	5.41	3.15	4.31
Minimum	5.74	-0.17	5.47	-1.73	4.17	4.05	2.28	2.83
Std. Dev.	0.06	0.11	0.17	0.17	0.22	0.22	0.19	0.23
Skewness	0.51	-1.65	-0.88	-2.95	-1.11	-1.73	0.82	-1.61
Kurtosis	4.70	8.56	4.39	15.63	4.48	8.02	3.04	7.34
Jarque-Bera	45.31	482.88	58.03	2242.56	82.30	430.03	31.35	336.61
Probability	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
BDS(1) test	***	***	***	***	***	***	***	***

Table 6. Descriptive statistics.
Notes: Descriptive statistics for log-levels data. Period: January 2015 to May 2020 - 277 observations. (1) BDS test stands for Broock, Scheinkman, Dechert, and LeBaron (1996) to determine whether the series were defined by nonlinearities. *** denote statistical significance at 1%.

Figure 5 shows the correlations between the log-transformed variables. In the preliminary analysis, the highest positive correlations were observed between WTI, HO, RBOB, and gasoline prices, suggesting close relationships in the energy markets. The ethanol price was also positively correlated with sugar and gasoline prices, highlighting the link between the prices of Brazilian fuels.

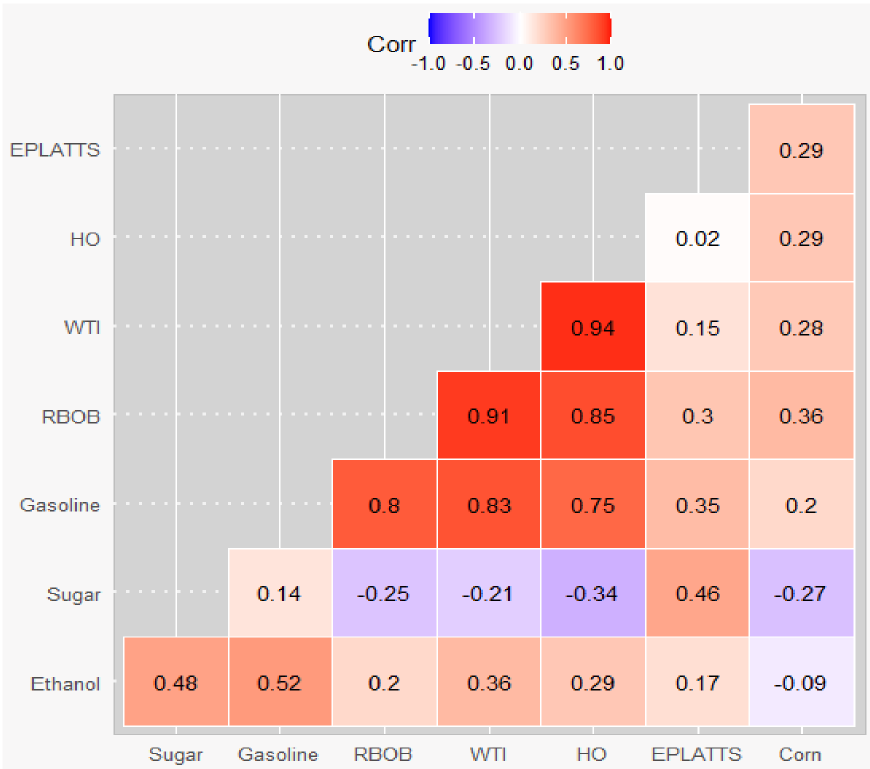


Figure 5. Pearson correlation coefficients for the relationships between log-transformed prices of energy and agricultural commodities.

Sugar held negative associations with the oil prices (RBOB, WTI and HO) mainly between 2017 and 2018 when India surpassed Brazil as the world's largest sugar producer. Thus, the world sugar surplus reduced the prices in the future market. While sugar prices remained low, oil prices went in the opposite direction. WTI hit more than $70 a barrel in mid-2018. Unsurprisingly, ethanol in Brazil achieved an all-time high predominance in the sugar mix (65% ethanol and 35% sugar). This supported the negative relationship with oil prices.

3.3.2
Econometric Approach

We investigated the symmetric and asymmetric effects on ethanol and gasoline prices in Brazil of the primary internationally traded energy commodities (WTI, RBOB, Ethanol Platts and HO) and the main feedstocks used in ethanol production (sugar and corn). We employed the NARDL framework proposed by Shin, Yu, & Greenwood-Nimmo (2014), which captures both nonlinear dynamics of long-term relationships and short-term dynamic adjustments. The approach overcomes the prevalent models in the literature – the error correction model (ECM), threshold ECM, Markov-switching ECM and smooth transition ECM – in jointly modelling the asymmetries and cointegration dynamics in a single step. Its flexibility and simplicity allow some hypotheses to be relaxed, such as the need for all variables involved being first order (E. M. de Oliveira et al., 2020). In this regard, NARDL offers robust estimation results even with a mix of I(0) and I(1) variables. Furthermore, the model allows accurate discernment of linear cointegration, nonlinear cointegration or a lack of cointegration (Katrakilidis & Trachanas, 2012).

We performed unit root tests of the log levels and first log differences of the data series in Table 7. We applied the following tests: augmented Dickey-Fuller (ADF) test (Dickey & Fuller, 1979); the PP test (Phillips & Perron, 1988); the ERS Point Optimal test (Elliott et al., 1996); the KPSS test (Kwiatkowski et al., 1992); and the KSS test of Kapetanios et al. (2003), which can detect nonstationarity even when the series contains significant nonlinear components.

Test type	Variable	ADF	PP	ERS	KPSS	KSS1	KSS2	KSS3
Log-levels								
Intercept	Corn	-3.69***	-3.98***	2.13**	0.14	-0.44	-3.10**	-3.10
	EPlatts	-3.21**	-2.53	1.94***	0.94***	-1.53	-2.95**	-3.58**
	Ethanol	-2.12	-2.08	4.89	0.29	-0.83	-2.62	-2.23
	Gasoline	-0.46	-0.01	13.69	0.23	0.77	-0.17	-0.29
	HO	-0.95	-1.08	6.41	0.48**	-0.78	-4.05***	-3.49**
	RBOB	-2.73	-2.47	2.45**	0.19	-0.56	-3.19**	-3.18*
	Sugar	-1.35	-1.60	7.04	0.60**	-0.80	-1.44	-1.65
	WTI	-1.13	-1.72	5.12	0.37*	-0.86	-2.41	2.11
First (log) differences								
Intercept	Δ ln Corn	-17.70***	-17.67***	0.63***	0.05	0.57	0.55	0.58
	Δ ln EPlatts	-8.50***	-14.99***	0.01***	0.10	-6.19***	-6.20***	-6.20***
	Δ ln Ethanol	-12.39***	-12.38***	0.24***	0.14	-6.07***	-6.14***	-6.31***
	Δ ln Gasoline	-10.96***	-11.18***	0.47***	0.26	-0.17	-0.25	-0.37
	Δ ln HO	-15.04***	-15.02***	0.31***	0.26	-8.14***	-8.13***	-8.05***
	Δ ln RBOB	-13.01***	-12.96***	0.48***	0.12	1.58	1.57	1.59
	Δ ln Sugar	-15.07***	-15.07***	0.26***	0.10	-4.49***	-4.48***	-4.53***
	Δ ln WTI	-12.18***	-12.53***	0.36***	0.20	-2.24	-2.28	-2.53

Table 7. Unit root tests for variables in log-levels and first log-differences
Notes: The statistics are pseudo t ratios for ADF and PP tests, PTμ and PTτ for the ERS (Point Optimal) test, LM statistic for the KPSS test, and tNL statistics for the KSS tests. The optimal lag lengths for ADF tests were selected by Schwarz Bayesian Information Criterion (Schwarz, 1978), and the bandwidth for PP and KPSS tests was selected with Newey-West using Bartlett kernel. ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively. KSS1, KSS2, and KSS3 correspond to the first (original data), second (demeaned), and third (demeaned and detrended) cases depicted in Kapetanios, Shin, and Snell (2003).

Overall, the results suggest that most series are integrated of the first order. The only cases in which we rejected the null hypothesis of the unit root test was for the corn and ethanol EPlatts. During our time series sample, both commodities' prices hovered around the mean with a steady behaviour, even though prices plunged during the COVID-19 pandemic. Nevertheless, as outlined, the NARDL allows for different orders of integration among the involved variables.

In the NARDL framework, the linear error correction mechanism (ECM) ((Engle & Granger, 1987) is extended to account for both the short-and long-run asymmetries. In this regard, a bivariate model is expressed in the following form:

$$\Delta Y_t = \mu + \rho\, Y_{t-1} + \theta_X^+\, X_{t-1}^+ + \theta_X^-\, X_{t-1}^- + \sum_{i=0}^{p} \alpha_i\, \Delta Y_{t-i} + \sum_{i=0}^{q_p} \left(\omega_{i,X}^+\, \Delta X_{t-i}^+\right) + \sum_{i=0}^{q_n} \left(\omega_{i,X}^-\, \Delta X_{t-i}^-\right) \qquad (3)$$

where Y_t and X_t are the dependent and independent variables, respectively, and Δ stands for the first differences operator. The long-term asymmetric coefficients are expressed as $L_{X^+} = \theta_X^+/\rho$ and $L_{X^-} = \theta_X^-/\rho$. p and q indicate the lag orders for the dependent and independent variables, respectively, in the distributed lag part. A long-run asymmetric response of y can be tested by rejecting the following Wald test null hypothesis: $\theta_X^+ = \theta_X^-$. Differences between positive and negative short-run adjustments, in turn, are suggested when the following null hypothesis is rejected: $\sum_{i=0}^{q_p} \omega_{i,X}^+ = \sum_{i=0}^{q_n} \omega_{i,X}^-$.

We employ the diagnostic BDM t-test (Banerjee, Dolado, & Mestre, 1998) to account for the existence of cointegration by rejecting the null hypothesis of $\rho = 0$ against $\rho < 0$. Moreover, the significance of the model was tested by rejecting the joint null hypothesis of $\rho = \theta_X^+ = \theta_X^- = 0$ when applying the PSS F-test (Pesaran, Smith, & Shin, 2001).

3.3.3
Bivariate and Multivariate Analysis

To achieve the best-fit NARDL specification, we first performed bivariate Wald tests to detect the possible existence of short- and long-run asymmetries between the variables involved. The results are shown in Table 8. They suggest that RBOB, HO, and sugar had both short- and long-run asymmetric effects on Brazilian gasoline prices (see Appendix A for the cumulative dynamic multipliers). In the case of ethanol as a dependent variable, we noted that the same independent variables had the same asymmetric effects, except for HO in the long-run.

Dependent variable	Independent variable	W_{SR}	W_{LR}	Conclusion
Gasoline prices as dependent				
Gasoline	RBOB	4.62**	7.38***	SR and LR asymmetry
	WTI	0.02	0.01	SR and LR symmetry
	Heating Oil	9.27***	3.32*	SR and LR asymmetry
	EPlatts	2.57	0.05	SR and LR symmetry
	Sugar	42.31***	12.47***	SR and LR asymmetry
	Corn	0.04	0.14	SR and LR symmetry
Ethanol prices as dependent				
Ethanol	RBOB	9.99***	14.28***	SR and LR asymmetry
	WTI	0.27	0.27	SR and LR symmetry
	Heating Oil	5.56**	2.54	SR asymmetry and LR symmetry
	EPlatts	0.51	0.29	SR and LR symmetry
	Sugar	17.90***	19.38***	SR and LR asymmetry
	Corn	0.00	0.00	SR and LR symmetry

Table 8. Bivariate NARDL models - Results of the short- and long-run symmetry tests
Notes: All variables are log-transformed. W_{SR} and W_{LR} are the Wald tests for short and long-run asymmetry, respectively. ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Bivariate models can be generalised into a multivariate setting to investigate the outcomes for the prices of gasoline and ethanol in Brazil when these variables are affected by the simultaneous interaction of all listed independent variables in Table 8. In our framework, we initially considered all independent variables as asymmetric, as shown in equations (4) and (5):

$$
\begin{aligned}
&\Delta \ln(\boldsymbol{Gasoline})_t = \mu + \rho \ln(Gasoline)_{t-1} + \theta^+_{RBOB} \ln(RBOB)^+_{t-1} + \\
&\theta^-_{RBOB} \ln(RBOB)^-_{t-1} + \theta^+_{HO} \ln(HO)^+_{t-1} + \theta^-_{HO} \ln(HO)^-_{t-1} + \theta^+_{Sugar} \ln(Sugar)^+_{t-1} + \\
&\theta^-_{Sugar} \ln(Sugar)^-_{t-1} + \theta^+_{WTI} \ln(WTI)^+_{t-1} + \theta^-_{WTI} \ln(WTI)^-_{t-1} + \\
&\theta^+_{EPlatts} \ln(EPlatts)^+_{t-1} + \theta^-_{EPlatts} \ln(EPlatts)^-_{t-1} + \theta^+_{Corn} \ln(Corn)^+_{t-1} + \\
&\theta^-_{Corn} \ln(Corn)^-_{t-1} + \\
&\textstyle\sum_{i=0}^{p} \alpha_i \,\Delta \ln(Gasoline)_{t-i} + \\
&\textstyle\sum_{i=0}^{q_p}\left(\omega^+_{i,RBOB}\,\Delta \ln(RBOB)^+_{t-i}\right) + \sum_{i=0}^{q_n}\left(\omega^-_{i,RBOB}\,\Delta \ln(RBOB)^-_{t-i}\right) + \\
&\textstyle\sum_{i=0}^{r_p}\left(\omega^+_{i,HO}\,\Delta \ln(HO)^+_{t-i}\right) + \sum_{i=0}^{r_n}\left(\omega^-_{i,HO}\,\Delta \ln(HO)^-_{t-i}\right) + \\
&\textstyle\sum_{i=0}^{s_p}\left(\omega^+_{i,Sugar}\,\Delta \ln(Sugar)^+_{t-i}\right) + \sum_{i=0}^{s_n}\left(\omega^-_{i,Sugar}\,\Delta \ln(Sugar)^-_{t-i}\right) + \\
&\textstyle\sum_{i=0}^{t_p}\left(\omega^+_{i,WTI}\,\Delta \ln(WTI)^+_{t-i}\right) + \sum_{i=0}^{t_n}\left(\omega^-_{i,WTI}\,\Delta \ln(WTI)^-_{t-i}\right) + \\
&\textstyle\sum_{i=0}^{t_p}\left(\omega^+_{i,EPlatts}\,\Delta \ln(EPlatts)^+_{t-i}\right) + \sum_{i=0}^{t_n}\left(\omega^-_{i,EPlatts}\,\Delta \ln(EPlatts)^-_{t-i}\right) + \\
&\textstyle\sum_{i=0}^{u_p}\left(\omega^+_{i,Corn}\,\Delta \ln(Corn)^+_{t-i}\right) + \sum_{i=0}^{u_n}\left(\omega^-_{i,Corn}\,\Delta \ln(Corn)^-_{t-i}\right) \qquad (4)
\end{aligned}
$$

$$\Delta \ln(\boldsymbol{Ethanol})_t = \mu + \rho \ln(Ethanol)_{t-1} + \theta_{RBOB}^{+} \ln(RBOB)_{t-1}^{+} + \theta_{RBOB}^{-} \ln(RBOB)_{t-1}^{-} + \theta_{HO}^{+} \ln(HO)_{t-1}^{+} + \theta_{HO}^{-} \ln(HO)_{t-1}^{-} + \theta_{Sugar}^{+} \ln(Sugar)_{t-1}^{+} + \theta_{Sugar}^{-} \ln(Sugar)_{t-1}^{-} + \theta_{WTI}^{+} \ln(WTI)_{t-1}^{+} + \theta_{WTI}^{-} \ln(WTI)_{t-1}^{-} + \theta_{EPlatts}^{+} \ln(EPlatts)_{t-1}^{+} + \theta_{EPlatts}^{-} \ln(EPlatts)_{t-1}^{-} + \theta_{Corn}^{+} \ln(Corn)_{t-1}^{+} + \theta_{Corn}^{-} \ln(Corn)_{t-1}^{-} + \sum_{i=0}^{p} \alpha_i \Delta \ln(Ethanol)_{t-i} + \sum_{i=0}^{q_p}\left(\omega_{i,RBOB}^{+} \Delta \ln(RBOB)_{t-i}^{+}\right) + \sum_{i=0}^{q_n}\left(\omega_{i,RBOB}^{-} \Delta \ln(RBOB)_{t-i}^{-}\right) + \sum_{i=0}^{r_p}\left(\omega_{i,HO}^{+} \Delta \ln(HO)_{t-i}^{+}\right) + \sum_{i=0}^{r_n}\left(\omega_{i,HO}^{-} \Delta \ln(HO)_{t-i}^{-}\right) + \sum_{i=0}^{s_p}\left(\omega_{i,Sugar}^{+} \Delta \ln(Sugar)_{t-i}^{+}\right) + \sum_{i=0}^{s_n}\left(\omega_{i,Sugar}^{-} \Delta \ln(Sugar)_{t-i}^{-}\right) + \sum_{i=0}^{t_p}\left(\omega_{i,WTI}^{+} \Delta \ln(WTI)_{t-i}^{+}\right) + \sum_{i=0}^{t_n}\left(\omega_{i,WTI}^{-} \Delta \ln(WTI)_{t-i}^{-}\right) + \sum_{i=0}^{t_p}\left(\omega_{i,EPlatts}^{+} \Delta \ln(EPlatts)_{t-i}^{+}\right) + \sum_{i=0}^{t_n}\left(\omega_{i,EPlatts}^{-} \Delta \ln(EPlatts)_{t-i}^{-}\right) + \sum_{i=0}^{u_p}\left(\omega_{i,Corn}^{+} \Delta \ln(Corn)_{t-i}^{+}\right) + \sum_{i=0}^{u_n}\left(\omega_{i,Corn}^{-} \Delta \ln(Corn)_{t-i}^{-}\right) \quad (5)$$

Equations (4) and (5) show the significant sources of price transmission from the international futures markets to Brazilian fuel prices. The positive and negative long-run transmission coefficients were denoted by θ^{+} and θ^{-}, while the positive and negative short-run transmission coefficients were represented by $\omega^{+}\ and\ \omega^{-}$. We initially considered that all independent variables had an asymmetric effect on the prices of gasoline and ethanol, both in the short run and long run. Then, depending on the symmetry or asymmetry of each coefficient based on Wald tests, we reduced the model's complexity by changing the short- and/or long-run effects of specific variables into being symmetric. The model selection also considered the values of specific log-likelihood criteria, namely the Akaike Information Criterion (AIC; Akaike, 1974) and the Schwarz Bayesian Information Criterion (BIC; Schwarz, 1978). In addition, the residuals from the selected models were free from serial correlation issues. Finally, whenever feasible, we also made sure that there was no heteroscedasticity in the residuals.

3.3.4
Nonlinear causality tests (Diks & Panchenko, 2006)

We extend our analysis by assessing the nonlinear causal relationships between the log-returns of fuel prices and energy/agricultural commodities prices returns. To this end, we apply the Diks and Panchenko (2006) (DP hereinafter) nonparametric test for Granger non-causality.

According to (1991), nonlinear models are better at capturing the behaviour of the financial market, such as sudden bursts of volatility and occasionally large movements. The seminal work on causality, proposed by Granger (1969), assumes a parametric and linear relationship for the conditional mean, which fails to detect nonlinear relationships over time. These include asymmetry, persistence and structural breaks (Baek and Brock,1992). In that regard, Baek et al. (1992) showed that nonlinear models are more suitable for detecting structural breaks and persistence in the variables, and they proposed a nonparametric test. Hiemstra et al. (1994) modified the Baek and Brock nonparametric causality test by relaxing the hypothesis that the time series is independent and identically distributed to test the dynamic relationships. Further, DP improved the Hiemstra and Jones test to overcome the possibility of spurious rejections of the null hypothesis.

The non-causality test aims to identify evidence against the null hypothesis

$$H_0: \{X_t\}\ is\ not\ Granger\ causing\ \{Y_t\}.$$

Under the null hypothesis, past observations of $\{X_t\}$ do not contain additional information on Y_{t+1}. DP consider $Z_t = Y_{t+1}$ and remove the time indices *t*. Thus, X and Z are conditionally independent in $Y = y$ for each fixed value of y. DP reformulated in terms of joint distribution, the conditional distribution of Z given (X, Y) = (x, y) equals to Z given Y=y. The joint probability density function $f_{X,Y,Z}(x, y, z)$ and marginal distributions must satisfy:

$$H_0: \frac{f_{X,Y,Z}(x,y,z)}{f_Y(y)} = \frac{f_{X,Y}(x,y)}{f_Y(y)} \cdot \frac{f_{Y,Z}(y,z)}{f_Y(y)} \tag{6}$$

DP shows another way of writing the null hypothesis of nonlinear Granger causality:

$$H_0: q \equiv E\big[f_{X,Y,Z}(X,Y,Z)\, f_Y(Y) - f_{X,Y}(X,Y)\, f_{Y,Z}(Y,Z)\big] = 0 \tag{7}$$

Let $\hat{f}_W(W_i)$ be a local density estimate of a random vector d_W - variate $\boldsymbol{W}$ in $\boldsymbol{W_i}$, defined by $\hat{f}_W(W_i) = \frac{(2\varepsilon_n)^{-d_W}}{n-1} \sum_{j j \neq i} I_{ij}^W$, where $I_{ij}^W = I\big(\|W_i - W_j\| < \varepsilon_n\big)$, $I(\cdot)$ is the indicator function, and ε_n is the bandwidth parameter dependent on n (number of observations); the statistical test is:

$$T_n(\varepsilon_n) = \frac{n-1}{n(n-2)} \sum_i \left(\hat{f}_{X,Z,Y}(X_i, Z_i, Y_i)\hat{f}_Y(Y_i) - \hat{f}_{X,Y}(X_i, Y_i)\hat{f}_{Y,Z}(Y_i, Z_i)\right) \tag{8}$$

Asymptotic properties are developed under the "mixing" conditions (Denker & Keller, 1983). When $\varepsilon_n = Cn^{-\beta}$, where C is a positive constant and $\beta \in (1/4, 1/3)$, we obtain:

$$\sqrt{n}\frac{(T_n(\varepsilon_n) - q)}{S_n} \xrightarrow{D} N(0,1) \tag{9}$$

where $\xrightarrow{D}$ denotes convergence in the distribution, and S_n is the estimator of the asymptotic variance of T_n. We set the bandwidth at 1($\varepsilon_n = 1$) according to some studies (Andreasson et al., 2016; E. M. de Oliveira et al., 2017).

		Lags				Lags	
X→	Y	*1*	*2*	*Y*→	X	*1*	*2*
				Ethanol			
Corn→	**Ethanol**	0.37	0.55	→	Corn	1.14	1.52
Eplatts→		-2.62	-2.08		Eplatts	0.91	-0.45
HO→		1.72**	2.20**		HO	0.8	-1.07
RBOB→		0.32	1.24		RBOB	1.06	0.58
Sugar→		0.70	0.43		Sugar	1.34	1.28
WTI→		1.29	1.46		WTI	1.3	0.27

		Lags				Lags	
X→	Y	*1*	*2*	*Y*→	X	*1*	*2*
				Gasoline			
Corn→	**Gasoline**	-0.26	-0.94	→	Corn	-0.56	-0.69
Eplatts→		-0.44	-2.09		Eplatts	0.67	-1.45
HO→		3.25***	1.72**		HO	0.91	-1.08
RBOB→		2.27**	1.82**		RBOB	1.03	-0.66
Sugar→		1.18	-0.40		Sugar	095	0.43
WTI→		2.40***	0.94		WTI	1.22	-0.13

Table 9. Nonlinear Causality Test (DP, 2006)

Notes: Nonlinear causality DP (2006) T_n statistic. (X) → (Y) denotes the independent variable does not cause the dependent variable and vice-versa. **, and *** show statistical significance at 5% and 1%, respectively. Bandwidth set at ε_n= 1 and lags ℓ=1,2.

The results show a unidirectional causality from HO to ethanol. This implies that HO is the main underlying commodity that affects the price of ethanol. The oil-ethanol nexus is likely to be stronger than the linkage between sugar and ethanol. Regarding the transmission to gasoline prices, the nonlinear test shows that RBOB and HO affect the price return of gasoline. The next section confirms whether these variables have an asymmetric effect on ethanol and gasoline prices.

Until 2015, we did not expect a transmission from global markets to the Brazilian fuel market since the government influenced gasoline prices as its discretion. Thus, we employed an additional nonlinear causality test from 2009 to 2014 to ensure that HO influence occurred mainly after 2015. The results[5] showed no causality running from the global commodities to gasoline or ethanol, corroborating our assumption.

[5] For conciseness, the results are not reported here. However, they are available upon request.

3.4
Results and Discussions

3.4.1
Main Results

The results from the selected multivariate NARDL models for Brazilian ethanol and gasoline prices are shown in Tables 10 and 11, respectively. HO had an asymmetrical effect on the ethanol price. The long-run positive coefficient, which was statistically significant, suggested that an increase of 1% in HO prices led to a 1.26% decrease in the price of ethanol in Brazil. Conversely, a 1% decrease in HO caused a 0.40% decrease in the price of ethanol. This was a much smaller effect, in both magnitude and statistical significance, than that from a positive shock. In the short run, declines in HO prices had a negative and statistically significant coefficient. The ethanol response exposed the mechanisms through which the Brazilian ethanol market adjusts its prices when international oil prices increase.

ΔEthanol as dependent				
Variable	Coefficient	Std. Error	Long-run	Coefficient
$Ethanol_{t-1}$	-0.11^{***}	0.03	-	-
HO^{+}_{t-1}	-0.14^{**}	0.07	L_{HO+}	-1.26
HO^{-}_{t-1}	0.04	0.07	L_{HO-}	0.40
$EPlatts_{t-1}$	0.07	0.04	$L_{EPlatts}$	0.59
$RBOB_{t-1}$	-0.12^{***}	0.04	L_{RBOB}	-1.06
WTI_{t-1}	0.16^{***}	0.06	L_{WTI}	1.41
$Sugar_{t-1}$	0.03	0.03	L_{Sugar+}	0.23
$Corn_{t-1}$	-0.13^{***}	0.05	L_{Corn}	-1.20
$\Delta Ethanol_{t-1}$	0.25^{***}	0.06		
ΔHO^{+}_{t-1}	-0.14	0.09		
ΔHO^{+}_{t-2}	0.08	0.07	Statistics	
ΔHO^{-}_{t-1}	-0.20^{**}	0.09	R^2	0.31
$\Delta EPlatts_{t-1}$	-0.10	0.08	AIC	-3.59
$\Delta EPlatts_{t-2}$	0.17^{**}	0.08	BIC	-3.34
$\Delta EPlatts_{t-3}$	0.12	0.08	χ^2_{FF}	1.40
$\Delta RBOB_{t-1}$	0.12^{*}	0.07	BG	2.54
ΔWTI_{t-1}	0.19^{**}	0.08	ARCH	10.17
$\Delta Corn_{t-2}$	0.08	0.09		
Constant	1.39^{***}	0.37		

Table 10. Multivariate NARDL - estimation results for the Brazilian Ethanol prices

Notes: All variables are log-transformed. First differences are denoted by the Δ operator and are computed on log-transformed data. $L_X = -\rho_X/\rho_{Ethanol}$ indicates the long-run effect of a specific regressor (X) on the Ethanol price. ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively. AIC and BIC stand for the Akaike Information Criterion (1974) and Schwarz Bayesian Information Criterion (1978), respectively. χ^2_{FF} is the Ramsey (1969) RESET / Functional Form Likelihood Ratio (LR) statistic. BG and ARCH refer to the empirical statistics of the Breusch–Godfrey Lagrange Multiplier (LM) test for serial correlation and the Engle (1982) test for conditional heteroscedasticity, both applied to residuals with 6 lags.

ΔGasoline as dependent				
Variable	Coefficient	Std. Error	Long-run	Coefficient
$Gasoline_{t-1}$	-0.07^{*}	0.04	-	-
$RBOB^{+}_{t-1}$	-0.07^{***}	0.04	L_{RBOB+}	-1.50
$RBOB^{-}_{t-1}$	-0.10	0.04	L_{RBOB-}	-0.47
WTI_{t-1}	-0.03^{*}	0.04	L_{WTI}	1.21
HO_{t-1}	0.08	0.04	L_{HO}	-0.36
$EPlatts_{t-1}$	-0.02^{*}	0.04	$L_{EPlatts}$	0.89
$Sugar_{-1}$	0.06^{**}	0.03	L_{Sugar}	-0.69
$Corn_{t-1}$	-0.05	0.02	L_{Corn}	-0.75
$\Delta Gasoline_{t-1}$	0.10	0.07		
$\Delta RBOB^{+}_{t-1}$	0.15^{**}	0.06		
$\Delta RBOB^{+}_{t-3}$	0.13^{**}	0.06		
$\Delta RBOB^{+}_{t-5}$	0.13^{***}	0.04	Statistics	
$\Delta RBOB^{-}_{t-1}$	0.19^{***}	0.06	R^2	0.42
$\Delta RBOB^{-}_{t-2}$	0.12^{**}	0.06	AIC	-4.24
$\Delta RBOB^{-}_{t-5}$	0.06	0.04	BIC	-3.96
ΔWTI_{t-1}	0.09^{*}	0.05	χ^2_{FF}	0.95
ΔHO_{t-3}	-0.06	0.05	BG	7.47
ΔHO_{t-3}	0.04	0.04	ARCH	6.93
$\Delta Sugar_{t-2}$	0.07	0.05		
$\Delta Corn_{t-1}$	-0.05	0.06		
Constant	0.19	0.30		

Table 11. Multivariate NARDL - estimation results for the Brazilian Gasoline prices
Notes: $L_X = -\rho_X/\rho_{Gasoline}$ indicates the long-run effect of a specific regressor (X) on the Gasoline price. Other notes: see Table 10.

Surprisingly, the contributions of sugar to the price of ethanol had little impact, as had been suggested by the nonlinear causality test. It should be noted, however, that the channels through which commodity prices affect Brazilian biofuels differ considerably from the results of previous studies. For example, Fernandez-Perez, Frijns, and Tourani-Rad (2016) showed that corn and soybean prices (mostly used in biofuel production) affect ethanol (futures prices in CBOT), and crude oil has a direct impact on agricultural commodities. In addition, Vacha et al. (2013) found that ethanol prices are primarily connected to the price of its US feedstock, and biodiesel prices are strongly connected to German diesel prices. According to Kristoufek et al. (2014), price transmission between corn and ethanol was clearer during the food crisis (2007–2008). On the other hand, Balcombe and Rapsomanikis (2008) examined long-term relationships in the Brazilian sugar–ethanol–oil nexus, and they found evidence that oil prices dictate the sugar and ethanol long-run equilibrium and that the direction of the effect is from oil to sugar to ethanol. However, the relationship between the prices of oil and ethanol may remain static when supply shocks in the sugarcane sector result in high prices for sugar and ethanol. In a broader sense, depending on the spread between prices, the sugar–ethanol–oil nexus can have different price adjustments. For instance, in our study, the shock from HO, coupled with low international sugar prices and incentives for Brazilian ethanol production, may have intensified the relationship between oil and ethanol. The negative correlation between the prices of oil and sugar, mainly in 2017-2018, strengthened the ethanol-oil nexus.

On the one hand, our work shows that the main domestic factor that drives hydrous ethanol is gasoline. Melo et al. (2014) demonstrated that ethanol demand grows after increases in gasoline prices and that the gasoline price ceiling guides the ethanol price dynamics. Cavalcanti et al. (2012) supported this line of reasoning. They showed that gasoline prices had an asymmetric effect on ethanol from 2003 to 2009. Their findings suggested that a 1% increase in the price of gasohol led to a 2.02% decrease in ethanol price in the long run. However, in the short-run their results showed a positive asymmetric effect. Hence, flex-fuel car owners tended to first substitute gasohol for hydrous ethanol and then they switched back to gasohol as ethanol prices rose. On the other hand, our findings suggest there is a pass-through of international oil prices to ethanol, particularly the HO. Therefore, the main effects of oil prices on the global markets funnelled from the HO to the price of ethanol. One possible explanation was the influence of the distributors on price formation in the refineries.

Nonetheless, the ethanol market in Brazil has faced significant structural changes over the years, such as the increase in ethanol in the gasoline mixture (from E25 to E27 in 2015). This improved the substitutability power, the changes in price policy adopted by Petrobras in 2016, the recent urge for countries to reduce CO_2 emissions combined with the low-carbon initiatives of the RenovaBio program, and the global slump in energy prices because of the COVID-19 pandemic. These factors have contributed to a new scenario, and they put a question mark over the next steps for ethanol development in Brazil.

In the case of the long-term effects on Brazilian gasoline prices, as shown in Table 10, a 1% increase in the price for RBOB led to a 1.50% decrease in gasoline prices, and a 1% decrease in the former led to a 0.47% increase in the latter. Negative changes in RBOB prices, when compared to positive changes, had significantly fewer effects on Brazilian gasoline prices. However, in the short term, the surges in RBOB had a positive impact on gasoline prices. The inverse relationship between short- and long-term effects can be influenced by the crude oil price dynamics and the use of ethanol as a substitute fuel. Hammoudeh et al. (2003) pointed out that, despite the influence of WTI on gasoline prices in the US and international HO prices, there is a higher directional impact of HO on gasoline prices. Moreover, they showed that the differences in the magnitude of the negative relationship between HO and gasoline are linked to seasonal pattern fluctuations.

Our findings show that in the short term, global price changes can affect gasoline price dynamics in Brazil. This suggests that gasoline prices were more elastic towards international commodities prices. Moreover, the Diks and Panchenko (2006) nonlinear causality test also showed that oil prices (RBOB, HO and WTI) had a unidirectional impact on gasoline prices, while only HO affected ethanol. Our work exposes the fuel market idiosyncrasies in Brazil, such as the dependency on diesel importation and consumption and the distribution market concentration. Moreover, we reveal the intense government interference on domestic gasoline and diesel prices before 2015 but backsliding afterward. This demonstrates the interconnection between international commodities prices and domestic fuel prices, previously reported in nonlinear causality tests.

To determine robustness, we assessed the changes in the signs and magnitude of the coefficients in Tables 10 and 11 under different time horizons. The results of the rolling regression graphs depicted in Appendix B show that the coefficients do not change much across different samples. This demonstrates the robustness of the selected models.

3.4.2 Limitations

We showed that our models are stable, free from heteroscedasticity and autocorrelation, and we selected the lowest value of information criteria (BIC and AIC) during the formulation process (see Tables 10 and 11). We also showed in Appendix B that our results are consistent under different time horizons. Nonetheless, official ethanol and gasoline data for Brazil are made available only weekly, whereas price adjustments may occur on any day of the week and the ethanol prices are the weekly average. Moreover, we also acknowledge the non-synchronous disparities regarding different time zones and different weekly settlements in the futures prices and the ethanol and gasoline prices. These factors can lead to inconsistencies in the econometric results because significant time lags are omitted from the regression analysis (Bachmeier & Griffin, 2003; Chua et al., 2017). As highlighted in the literature review, despite Petrobras' guidance to adjust prices according to international markets, there have been several cases when the government intervened in oil surges.

3.5 Conclusions, Implications and Suggestions for Future Studies

This paper explored the impacts of futures prices for international energy and agricultural commodities on the prices for ethanol and gasoline in Brazil. The empirical investigation was performed using a multivariate nonlinear autoregressive distributed lag (NARDL) model to capture the possible symmetric and asymmetric effects of the independent variables on fuel prices in both the short term and long term. To the best of our knowledge, this study is the first to consider the impact of multiple variables on Brazil's fuel prices. Using weekly data from January 2015 to May 2020, our results showed an asymmetric effect of prices for reformulated blendstock for oxygenate blending (RBOB) on gasoline prices in Brazil in both the short and long run. The results also revealed that increases in heating oil (HO) prices led to declines in the price of ethanol in the long run, and this supported the long-term conversion from fossil fuels to ethanol. We also detected a nonlinear causality from HO to ethanol and RBOB to gasoline, which corroborates the asymmetric effects in the NARDL test.

Ethanol prices were shown to be more influenced by oil prices. On the other hand, sugar prices had a non-significant impact on ethanol prices in the long run. The nonlinear test showed no linkage between sugar and ethanol. Similar results were found by Balcombe and Rapsomanikis (2008) and (2012). Both works indicated that oil prices determined the long-run balance between the prices of sugar and ethanol in Brazil. Since the initial shock was generated by the price of heating oil, a similar shock may disrupt the sugar–ethanol nexus, reducing the direct impact of sugar on the price of ethanol. In addition, Allen et al. (2018) showed that ethanol as a dependent variable does not show a significant response to sugar shocks in low volatility regimes of the latter. Brazil's low international sugar prices coupled with ethanol incentives might have strengthened the oil–ethanol nexus.

The transmission mechanisms from international to domestic prices in the Brazilian fuel market have a wide range of implications for different stakeholders. First, policymakers may benefit from our results to see how to adjust price policies and hedging decisions to lessen the impacts of spikes in international oil prices and to extenuate the inflationary process. Countries that rely on oil imports and are sensitive to fluctuations in oil prices can apply the pass-through model to investigate price transmissions (see Sarwar et al., 2020).

Second, our results shed light on how domestic prices react to movements in international prices. Petrobras, Brazil's state-owned company and one of the world's largest oil producers, could take advantage of RBOB futures contracts as hedging devices for gasoline prices and improve its price policy design. In Brazil, there are no futures contracts for gasoline. Therefore, RBOB futures contracts could be an essential innovative hedging strategy. Likewise and most important, the ethanol sector could benefit from international futures contracts to improve the risk-sharing mechanism since market volatility has implications for input allocation, capital investments, and productivity growth. This may be useful not only to the Brazilian ethanol sector but for any country that trades ethanol.

Finally, institutions and investors who want to optimize their allocation strategy in the energy market need information about the complexity of the oil–sugar–ethanol nexus. Information about price transmission and asymmetric price adjustments is valuable to these financial actors, as they could adapt their short- and long-term trading strategies to improve their market positions.

One caveat of the present study is that we considered fuel prices based on the state of Sao Paulo, which represents the highest weight in average fuel consumption in Brazil, and where the largest domestic refinery is located. A potential avenue for future research would be to investigate disparities in local prices, logistics, and margins. Even though margins represent a small fraction of the final consumer price, further studies could help clarify the influence of the margin on gas station prices. In addition, it would be interesting to explore the effects of exchange rates on fuel prices, considering that Brazil's currency has had considerable depreciation since 2015, and investigate the impact of the COVID-19 pandemic situation to determine whether price transmission intensified or lessened. Further studies could also take advantage of the findings of this study to delve into the optimal hedging strategy using underlying futures contracts to offset the impact of volatility on domestic fuel prices.

3.6
Appendix. Bivariate NARDL models: cumulative dynamic multipliers

Figure A.1. shows the cumulative dynamic multipliers of the independent variables on ethanol prices, whilst Figure A.2 depicts the same multipliers when gasoline prices are considered as the dependent variable.

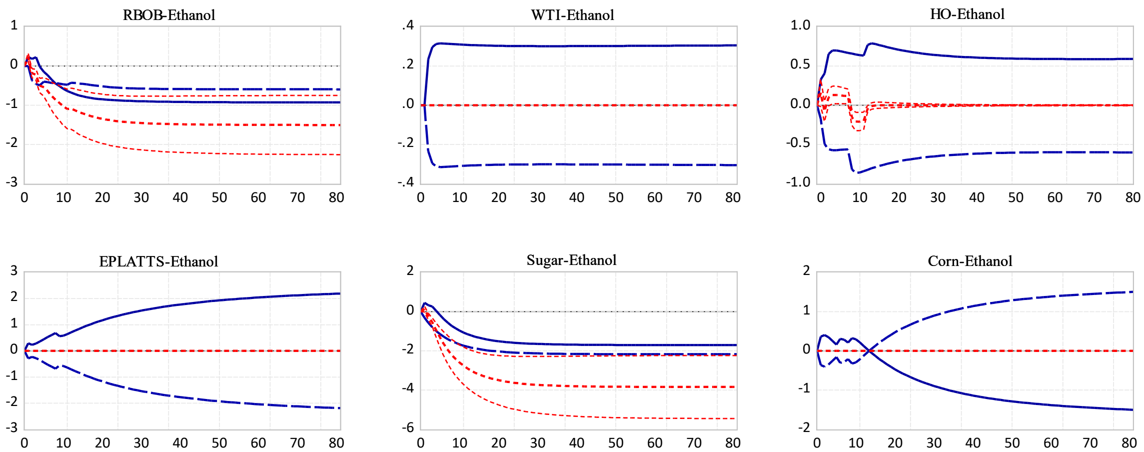


Figure A.1. Bivariate NARDL models—cumulative dynamic multipliers—with ethanol as the dependent variable.

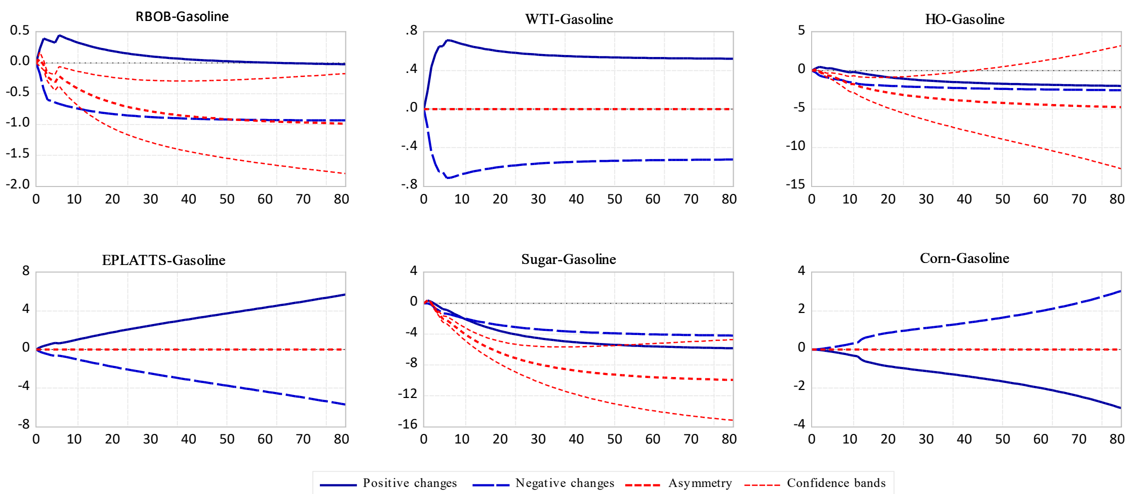
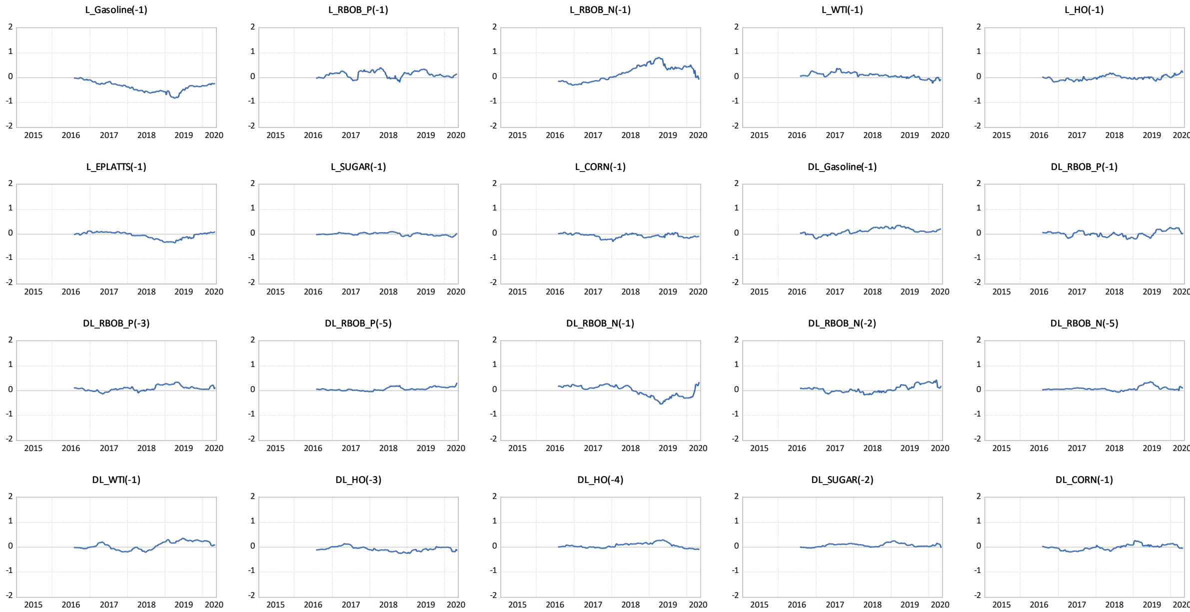


Figure A.2. Bivariate NARDL models—cumulative dynamic multipliers—with gasoline as the dependent variable.

3.7
Appendix. Rolling regressions: robustness check

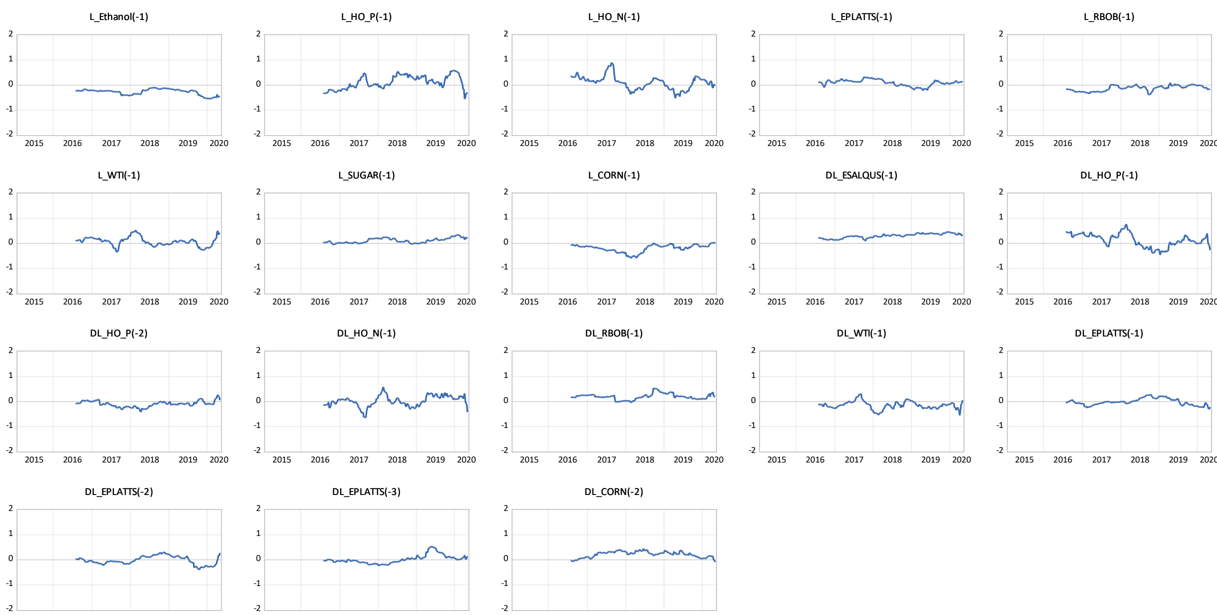
This appendix illustrates the results from a rolling regression equation using the same multivariate NARDL models selected for each dependent variable: gasoline and ethanol. We use a fixed window size of 80 weeks of observations and collected the estimated coefficients for each sliding step (1 week). A smaller window would have led to various issues related to the substantial number of coefficients to be estimated in the NARDL model. We generated models for 20 different periods by using a rolling window of 80 weeks.

The coefficients shown in the rolling regression graphs do not change much across different samples, which shows evidence of the robustness of the results presented in Section 4.



B.1. NARDL for gasoline prices – Rolling coefficients

Notes: D stands for first log differences and L for log-transformed variables. P and N correspond to positive and negative changes, respectively, in the underlying variable.



B.2. NARDL for ethanol prices: rolling coefficients

Notes: See notes in Figure B.1.

4
Volatility spillovers in the Brazilian agricultural markets.

4.1
Introduction

Understanding volatility spillover effects on asset prices is crucial to fully grasp financial markets' dynamic pricing information. Pricing information outlines additional inputs for the agents' decision-making process in an increasingly connected market. For example, volatility spillover information may prove useful for public policy recommendations, optimize trading strategies, and improve risk management effectiveness.

An analysis of the volatility spillover effect in agricultural commodity markets may provide additional information input for efficient resource allocation decisions about harvesting, output, storage, commercialization, and hedging. To improve these allocative decisions, large agricultural commodity players and policymakers may benefit from the volatility spillover analysis of prices and returns (Balcilar & Bekun, 2020; Fasanya & Odudu, 2020). Therefore, volatility spillover analysis may improve hedging strategies and risk management. Notably, agricultural commodity domestic futures markets show thin liquidity, indicating low hedging and portfolio diversification capacity.

We evaluate the Brazilian volatility spillovers in agricultural commodities between spot markets and futures markets. We apply the Diebold and Yilmaz (DY) (2009; 2012; 2014) framework for Brazilian coffee, ethanol, soybeans, and reformulated blendstock for oxygenate blending (RBOB) using spot and futures prices to compute the total volatility spillover, the directional spillover *to* and *from*, and the net pairwise volatility spillover.

The remainder of this paper is organized as follows. Section 2 outlines the literature on volatility spillover and hedging. Section 3 discusses the methods and data used this study, in particular, the Diebold and Yilmaz (DY) (2009; 2012; 2014) framework. Section 4 presents the results and a discussion of the main research findings. Last, Section 5 summarizes the results and suggests avenues for future research.

4.2
Literature review

The extant literature regarding volatility spillover is vast, particularly research done after the 2008 financial crisis. The Diebold and Yilmaz (DY) (2009; 2012; 2014) model is widely used to examine the volatility connectedness and the spillover effects among markets.

In 2009, Diebold and Yilmaz (2009) composed a measure of interdependence of asset returns and volatilities, examining the *return spillovers* and the *volatility spillovers*. The framework facilitates the analysis of non-crisis and crisis periods, including trends and bursts in the spillover effect. Their study uses data from 19 global equity markets from 1990 to 2009, showing divergent behavior in the dynamics of return spillovers versus volatility spillovers where return spillovers demonstrated a smooth increasing trend but no bursts.

Diebold and Yilmaz (2012) later propose a different approach using a generalized vector autoregressive framework in which forecast-error variance decompositions are invariant to the variable ordering. Diebold and Yilmaz examine daily volatility spillovers across US stock, bond, foreign exchange, and commodities markets, showing that significant volatility fluctuations in all four markets during the sample and cross-market volatility spillovers were limited until the global financial crisis began in 2007. Spillovers from the stock market to other markets increased after the Lehman Brothers' collapse in September 2008.

In another study, Diebold and Yılmaz (2014) also evaluate the connectedness measures composed from parts of variance decompositions, showing useful connectedness measures. They analyze the daily time-varying connectedness of major US financial institutions' stock return volatilities in recent years, including the financial crisis of 2007–2008.

Authors	**Period**	**Asset**	**Findings**
Chevallier and Ielpo (2013)	1995–2012	Standart assets, commodities, and currencies	Commodities exhibited weaker volatility spillover
Antonakakis and Kizys (2015)	1987-2014	Gold, silver, platinum, CHF/USD, and GBP/USD exchange rates	Gold, silver, platinum, CHF/USD, and GBP/USD were net transmitters of returns spillovers
Diebold, Liu, and Yilmaz (2017)	2011-2016	Nineteen commodities.	The energy sector was the highest contributor to other commodities. Energy, industrial metals, and precious metals were highly connected
Zhang and Broadstock (2020)	1982-2017	Beverage, fertilizer, food, metal, precious metals, raw materials, and oil	Food commodities contributed to the system dynamic more than 80% after 2008
Dahl, Oglend, and Yahya (2020)	1986-2016	Crude oil and agricultural commodities	Crude oil became net receiver after 2006, and during periods of financial turmoil, evidence of bidirectional spillover between crude oil and agricultural commodities.
Yoon et al. (2019)	1999-2016	Stock, bond, currency, and commodities	US stock market was the largest contributor of return spillover in the Asia-Pacific.
Balcilar and Bekun (2020)	2006-2016	Cocoa, banana, groundnut, soybeans, barley, maize, sorghum, rice, wheat, CPI, NOIL, and NEER.	Banana, cocoa, groundnut, maize, soybeans, and wheat were net transmitters of spillovers.
Fasanya and Odudu (2020)	1980-2017	Wheat, rice, soybeans, groundnut, and palm oil	Interdependence among agricultural commodities in Nigeria

Table 12. Summary of the literature on volatility spillover in commodities markets

4.3
Methodology and data

4.3.1
Spillover approach

We use the Diebold and Yilmaz (DY) (2009, 2012; 2014) framework to calculate the volatility spillover index among 12 Brazilian commodity spot markets and international commodity and financial futures markets. In particular, the index derives from the variance decomposition of an n-variable vector autoregression (VAR) model used to calculate the total spillovers in a simple VAR, with Cholesky factor orthogonalization.

Next, the DY approach generates directional spillovers in a generalized VAR framework, eliminating dependence on ordering results. As such, a covariance stationary n-variable VAR (p):

$$x_t = \sum_{i=1}^{p} \Phi_i x_{t-i} + \varepsilon_t \qquad \text{Eq. 1}$$

where: $\varepsilon \sim (0, \Sigma)$ = vector of i.i.d. disturbances.

The moving average representation is: $x_t = \sum_{i=0}^{\infty} A_i \varepsilon_{t-i}$, with the n x n coefficient matrices expressing $A_i = \Phi_1 A_{i-1} + \Phi_2 A_{i-2} + \cdots + \Phi_p A_{i-p}$, and A_0 = n x n identity matrix, with $A_i = 0$ for $i < 0$. Specifically, the DY approach calculates the variance decomposition, expressing each variable's forecast error variances in system shocks. Therefore, the variance decomposition demonstrates the fraction of the h-step-ahead error variance in forecasting x_i resulting from shocks to x_j, for all $j \neq i$, for each i.

The DY index uses the VAR framework of Koop et al. (1996) and Pesaran and Shin (1998) (KPSS), resulting in variance decompositions that are order invariant. As such, the shocks to each variable are not orthogonal, and the sum of the contributions to the forecast error variance, the row sum of the parts of the variance decomposition table, cannot be equal to one.

Next, we define the own variance shares as the fractions of the h-step-ahead error variances in forecasting x_i, resulting in shocks to x_i, for i = 1, 2, …, n, and cross variance shares, or spillovers, as the fractions of the h-step-ahead error variances in forecasting x_i relating to shocks to x_j, for i, j = 1, 2, …, n, such that $i \neq j$. Defining the KPSS h-step-ahead forecast error variance decompositions by $\theta_{ij}^{g}(h)$, for h = 1, 2, …, then:

$$\theta_{ij}^{g}(h) = \frac{\sigma_{ij}^{-1} \sum_{h=0}^{H-1} \left(e_i' A_h \Sigma \, e_j\right)^2}{\sum_{h=0}^{H-1} \left(e_i' A_h \Sigma A_h' e_j\right)} \qquad \text{Eq. 2}$$

where: Σ = variance matrix for the error vector e; σ_{ij} = standard deviation of the error term for the j-th equation; and, ε_i = selection vector, with one as the i-th element, and zeros otherwise.

To estimate the total volatility spillover index, we use the volatility contributions from the KPSS variance decomposition. The total spillover index demonstrates the contribution of spillovers of volatility shocks of the asset classes to the total forecast error variance:

$$S^g(h) = \frac{\sum_{\substack{i,j=1 \\ i \neq j}}^{n} \tilde{\theta}_{ij}^g(h)}{\sum_{i,j=1}^{n} \tilde{\theta}_{ij}^g(h)} \cdot 100 = \frac{\sum_{\substack{i,j=1 \\ i \neq j}}^{n} \tilde{\theta}_{ij}^g(h)}{n} \cdot 100 \quad \text{Eq. 3}$$

The DY framework estimates the net volatility spillover from market *i* to all other markets *j*, the difference between the gross volatility shocks transmitted to and received from all other markets, as:

$$S_i^g(h) = S_{i \to j}^g(h) - S_{i \leftarrow j}^g(h) \text{ Eq. 4}$$

Last, we compute the net pairwise volatility spillovers between markets *i* and *j*, the difference between the gross volatility shocks transmitted from market *i* to market *j*, and the shocks transmitted from *j* to *i*:

$$S_{ij}^g(h) = \left(\frac{\tilde{\theta}_{ji}^g(h)}{\sum_{i,k=1}^{n} \tilde{\theta}_{ik}^g(h)} - \frac{\tilde{\theta}_{ij}^g(h)}{\sum_{j,k=1}^{n} \tilde{\theta}_{jk}^g(h)} \right) \cdot 100$$

$$= \left(\frac{\tilde{\theta}_{ji}^g(h) - \tilde{\theta}_{ij}^g(h)}{n} \right) \cdot 100 \quad \text{Eq. 5}$$

To estimate the volatility values, we model the weekly log returns: $r_t^i = \log(\frac{P_t^i}{P_{t-1}^i})$, where r_t^i = weekly log return of commodity *i*, week *t*, and P_t^i = price of commodity *i*, week *t*. Then, we apply a standard *GARCH* (*1,1*) model on the weekly log returns, r_t^i, taking the square root of the resulting variance to calculate the weekly volatility of commodity *i* on week *t*, σ_t^i.

4.3.2
Data

We use weekly data for the following spot and futures prices series, encompassing three Brazilian commodity groups, softs, grains and oilseeds, and energy (see the descriptive statistics in Section 4.5) from September 2010 to March 2020.

We estimate the QR, MV, and ECM optimal hedge ratio for each group, using different contracts. For the softs group, we use the CEPEA/ESALQ Arabica Coffee as the spot price and the July 2020 ICE coffee weekly futures price series. For the energy group, we apply the CEPEA/ESALQ São Paulo Hydrous Ethanol Fuel Indicator as the spot price. For the future price, we use the May 2020 CME Ethanol futures prices and the May 2020 CME reformulated blendstock for oxygenate blending (RBOB) gasoline futures prices. We also compare both ethanol and RBOB futures optimal hedge ratios set—QR, MV, and ECM. For the grains and oilseeds group, we use the ESALQ/BM&FBOVESPA Paranaguá Soybeans spot prices, and for the futures prices, we use the March 2020 CME soybeans weekly prices. We use the March 2020 soybeans futures contracts due to its synchronicity with the Brazilian soybeans harvest season.

4.4
Empirical results and discussion

The first section presents the volatility spillover results for the Brazilian agricultural commodity markets—coffee, ethanol, soybeans, and RBOB—weekly log returns.

4.4.1
Spillover effects

Table 13 shows the results of gross directional spillovers. The results are based on VAR with max lags of 5 and generalized variance decomposition of 10^{th} day-ahead volatility forecast errors (Diebold & Yilmaz, 2012).

The sums of the rows represent the contributions 'FROM Others,' and the sums of the columns show the contribution ‘TO Others.’ The total volatility spillover index is 41.6% (lower right corner). The coffee futures prices express the highest contribution from others at 59.32%, followed by coffee futures at 46.52%, while the coffee RBOB shows the lowest contribution at 29.14%. Accounting for the net spillover, we find the largest spillover effect from the soybean futures prices (SOY_F: 82.51-38.70 = 43.81%) and from others to the ethanol spot price (ETH_S: 11.09-36.86 = -25.77%). Notably, there was an abrupt spike in 2020, highlighting the volatility spillover increase due to the COVID-19 pandemic (see Figure 6).

	ETH_F	**ETH_S**	**COFFEE_F**	**COFFEE_S**	**SOY_S**	**SOY_F**	**RBOB**	**FROM Others**
ETH_F	62.29	0.78	5.18	3.84	2.33	23.32	2.25	37.71
ETH_S	10.90	63.14	4.59	9.30	0.59	8.15	3.33	36.86
COFFEE_F	5.57	2.07	53.48	31.04	1.98	5.57	0.29	46.52
COFFEE_S	6.22	2.47	41.54	40.68	1.63	7.16	0.29	59.32
SOY_S	8.64	1.19	1.49	1.74	57.07	29.14	0.72	42.93
SOY_F	14.38	0.10	2.19	2.16	19.37	61.30	0.50	38.70
Rbob	9.16	4.48	0.81	2.50	3.04	9.16	70.86	29.14
Directional TO Others	54.87	11.09	55.80	50.59	28.94	82.51	7.37	
Directional Including Own	117.16	74.23	109.28	91.27	86.01	143.81	78.24	Total Spillover **41.6%**

Obs.: F = futures prices; S = spot prices. ETH stands for Ethanol and SOY for soybeans.

Table 13. Volatility spillover (connectedness) for weekly log returns Spillover (Connectedness)

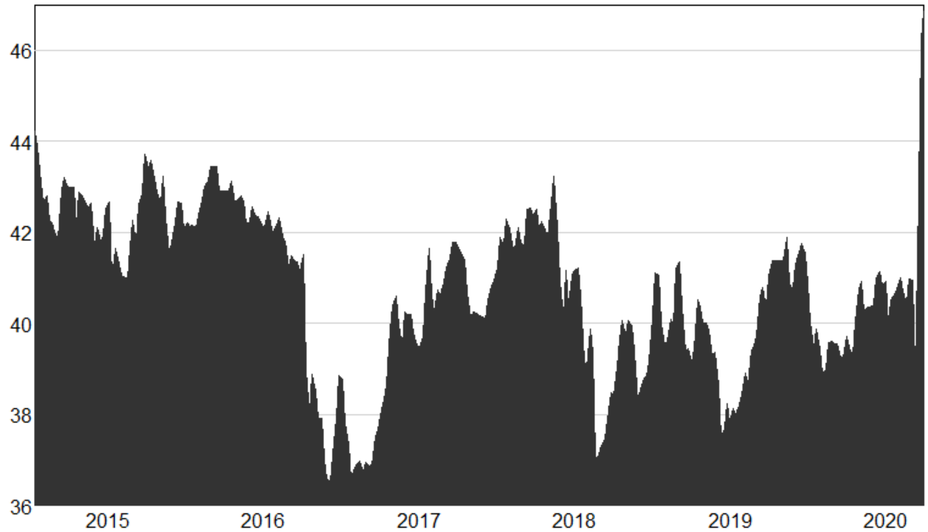


Figure 6. Total Volatility Spillover (Connectedness): coffee, ethanol, soybean, and RBOB.

The total spillover shows the volatility from all commodities analyzed in our study. Nonetheless, we are also interested in understanding the directional spillover information. Thus, we estimate and plot the $S^g_{i\leftarrow j}(h)$ directional 'TO Others' row, and the $S^g_{i\rightarrow j}(h)$ directional 'FROM Others' column. Figure 7 depicts the directional spillover to others; among the commodities, either soybean futures or spot are the greatest gross contributors to other markets. Figure 8 illustrates the directional volatility spillovers from others to all the involved variables. The spillover from others to coffee futures price appears to be increasing over time, while the coffee spot remains steady.

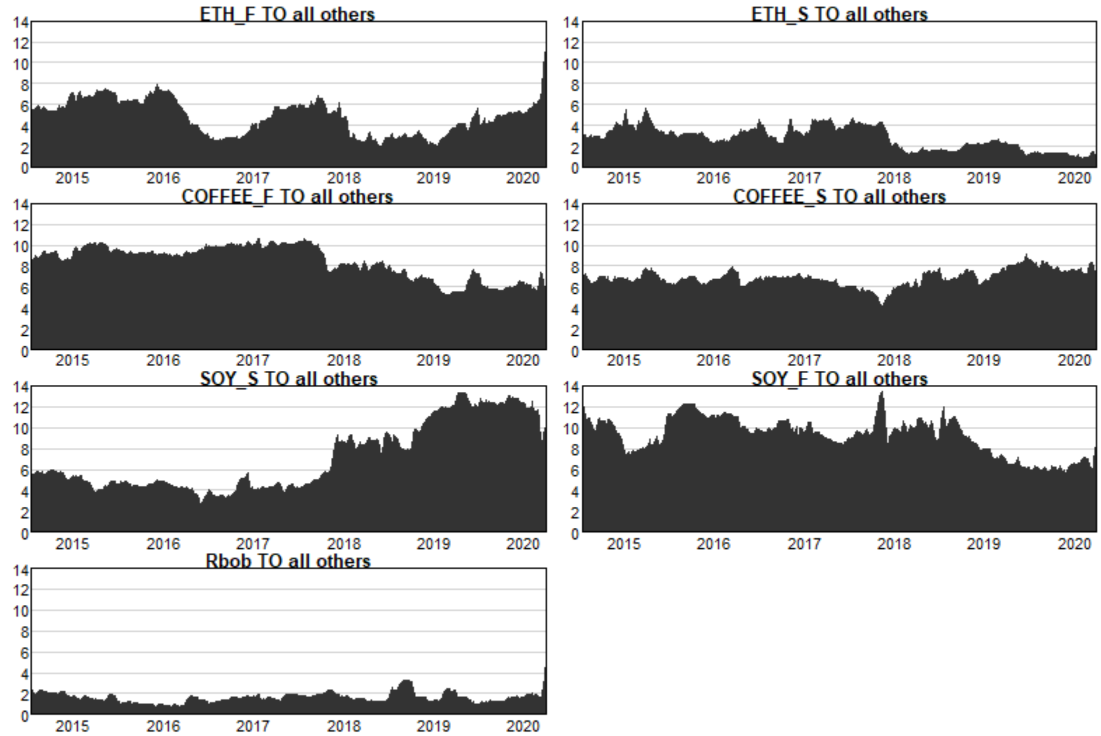


Figure 7. Directional Volatility Spillovers, TO ethanol (ETH), coffee, soybean (SOY), and RBOB

Obs.: F = futures prices; S = spot prices.

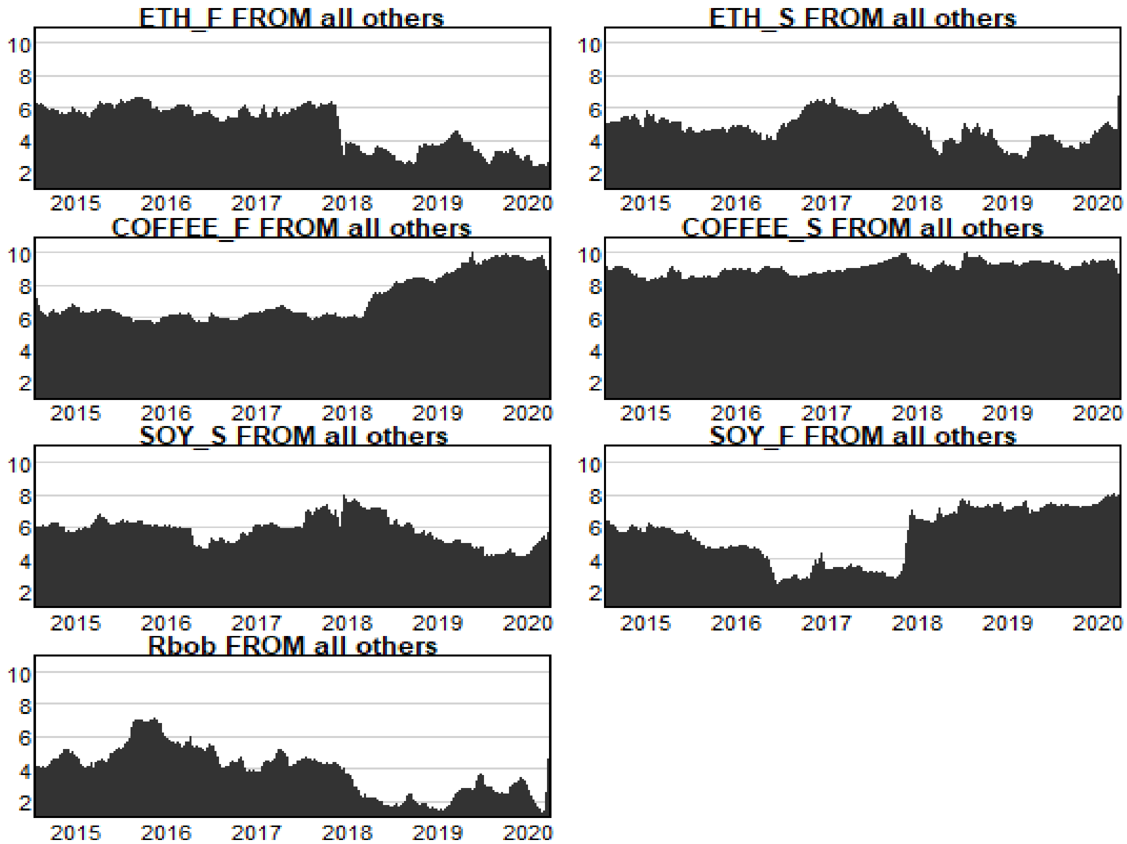


Figure 8. Directional Volatility Spillovers, FROM ethanol (ETH), coffee, soybean (SOY), and RBOB

Obs.: F = futures prices; S = spot price

We also provide the net pairwise combination of the involved variables (Equation 4). Thus, the plot in Appendix A.2. shows how much each commodity contributes to the volatility in other markets. The soybean spillover dynamic emphasizes the impacts of the trade war between China and the US, the increase of the Brazilian agricultural commodity exports, especially to China, and the commodity markets' financialization. For instance, the soybean spot became the net transmitter after 2018, coincidentally after the beginning of the trade war. On the other hand, the coffee futures also became net receivers after 2018, whereas the coffee spot maintains its net receiver pattern throughout our sample. Coffee appears to be more elastic in comparison to other commodities. Ethanol futures, in turn, were shown to be the net transmitter at the beginning of the pandemic outbreak, while ethanol spot prices remain steady during our sample analysis.

The directional volatility spillovers for the Brazilian agricultural commodity spot and futures markets show an idiosyncratic volatility pattern for each Brazilian commodity group. The DY framework indicates different volatility regimes, coupled with the price level analysis, can help decipher the market pricing dynamics and trends. Therefore, the DY volatility spillover effect provides a helpful approach to upgrade informational inputs about volatility patterns in domestic and international markets, formulating more efficient resource allocation decisions. In particular, a deeper knowledge of the volatility spillover dynamics for the Brazilian agricultural commodity spot markets can result in increased competitiveness and informed production, storage, commercialization, and hedging decisions.

4.5 Conclusion and further research

Our study analyzed the volatility spillovers between the Brazilian agricultural commodity spot prices and futures markets. We calculated the total volatility spillover, the directional spillover to and from commodities, and the net pairwise volatility spillover, applying the Diebold and Yilmaz (DY) (2009; 2012; 2014) framework. An in-depth analysis of volatility spillover provides useful information about the market dynamic.

The recent increase in total spillover effects reveals the impact of the COVID-19 pandemic and the volatility increase through time. Furthermore, the pairwise spillovers expose the effects of the rise in Brazilian agricultural commodity exports, especially soybean exports to China due to the trade war between China and the US.

The DY volatility spillover framework demonstrates the usefulness of estimating the dynamic relationships among the variables. The DY framework outline illustrates a systematic application to examine volatility spillover dynamics in the Brazilian agricultural commodity spot markets, both regional and national: for example, to analyze the soybeans market volatility spillover between Mato Grosso and Paraná.

Further research could expand the study of the volatility spillover effects in a multivariate framework, using a non-parametric technique. Understanding the price dynamics is paramount not only for policymakers to design policies to alleviate the spike in food prices, but also for farmers, consumers, and countries since high volatility could jeopardize agricultural investment, mainly when there is no mechanism to share risk (i.e., futures contracts).

4.6 Appendix

Weekly

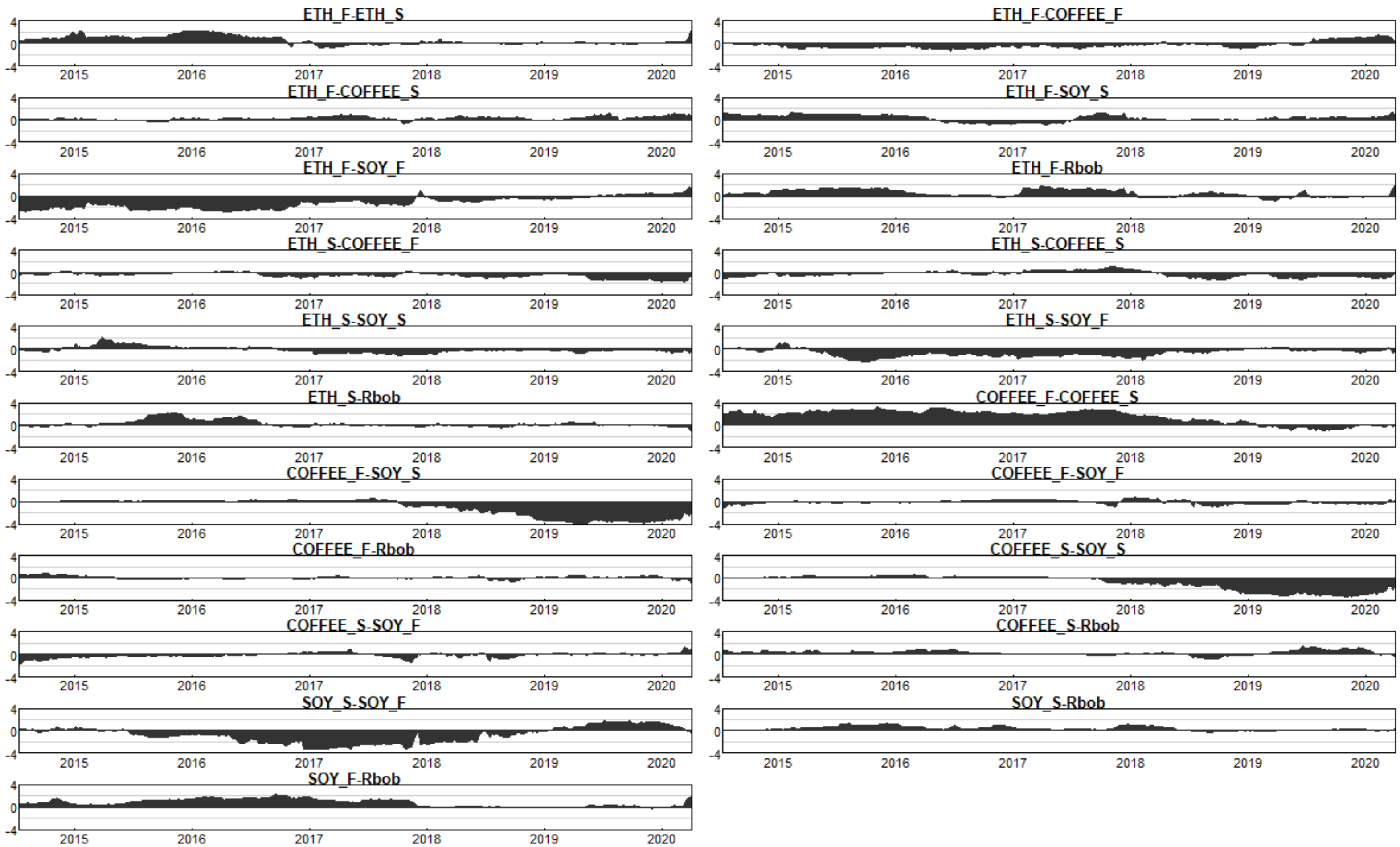
Statistics	Coffee$_F$	D(Coffee$_F$)	Coffee$_S$	D(Coffee$_S$)	Soy$_F$	D(Soy$_F$)	Soy$_S$	D(Soy$_S$)	ETH$_F$	D(ETH$_F$)	ETH$_S$	D(ETH$_S$)	RBOB$_F$	D(RBOB$_F$)
Mean	155.02	-0.14	162.79	-0.17	24.35	-0.01	26.19	-0.01	0.47	0.00	0.52	0.00	0.59	0.00
Median	139.13	-0.43	144.39	-0.63	22.66	0.02	24.16	0.01	0.41	0.00	0.50	0.00	0.54	0.00
Maximum	299.85	27.00	341.76	26.94	37.54	2.35	45.00	2.62	0.72	0.07	0.98	0.09	0.89	0.06
Minimum	89.00	-27.95	95.47	-21.71	18.55	-6.08	18.70	-6.81	0.23	-0.13	0.30	-0.19	0.16	-0.12
Std. Dev.	46.09	6.62	57.20	5.80	4.40	0.74	5.30	0.68	0.11	0.02	0.10	0.02	0.15	0.02
Skewness	1.14	0.22	1.40	0.37	0.66	-1.92	0.88	-2.11	0.76	-1.20	0.98	-1.69	0.14	-1.35
Kurtosis	3.53	5.09	4.26	5.53	2.27	16.77	3.33	24.87	2.15	11.45	4.53	18.32	1.83	8.21
JB	114	94	196	145	47	4239	67	10289	63	1599	128	5109	30	714
Prob	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ObS.	498	498	498	498	498	498	498	498	498	498	498	498	498	498

Monthly

Statistics	Coffee$_F$	D(Coffee$_F$)	Coffee$_S$	D(Coffee$_S$)	Soy$_F$	D(Soy$_F$)	Soy$_S$	D(Soy$_S$)	ETH$_F$	D(ETH$_F$)	ETH$_S$	D(ETH$_S$)	RBOB$_F$	D(RBOB$_F$)
Mean	214.72	-0.28	241.39	-0.20	29.79	0.06	31.80	0.06	0.62	0.00	0.64	0.00	0.74	0.00
Median	210.30	-0.25	242.50	-0.90	29.82	0.13	30.46	0.13	0.62	0.00	0.64	0.00	0.75	0.00
Maximum	299.85	24.75	341.76	17.91	37.54	2.35	45.00	2.62	0.72	0.07	0.98	0.09	0.89	0.06
Minimum	148.95	-27.95	161.93	-18.18	23.70	-2.72	25.27	-6.81	0.50	-0.08	0.49	-0.19	0.57	-0.10
Std. Dev.	42.07	8.13	55.07	7.50	2.73	0.92	4.57	1.00	0.05	0.02	0.10	0.03	0.07	0.03
Skewness	0.14	0.01	0.20	0.01	0.34	-0.40	1.10	-2.78	-0.21	-0.43	0.78	-2.11	-0.12	-1.30
Kurtosis	1.76	3.95	1.62	2.88	3.24	2.89	3.61	20.79	2.77	4.38	3.55	16.29	2.88	6.66
JB	8	5	11	0	3	3	27	1,781	1	13	14	997	0	103
Prob	0.02	0.10	0.01	0.96	0.27	0.19	-	-	0.56	-	-	-	0.83	-
ObS.	123	123	123	123	123	123	123	123	123	123	123	123	123	123

A. 1. Descriptive statistics. Coffee, ethanol, soybeans spot and futures prices, gasoline futures prices. Weekly and monthly values, levels, and first differences.

Source: Research results. Obs.: D(.) = first difference; F = futures prices; S = spot prices.



A.2. Net Pairwise Volatility Spillovers (ethanol, soybeans, coffee, and RBOB)

5
Can we measure ambiguity in commodity futures markets? Formulation of an empirical ambiguity measure

5.1 Introduction

Ambiguity is defined as the uncertainty in the probability distribution resulting from misinterpretation or lack of information. Ellsberg (1961) shows that under ambiguous payoffs, agents strictly prefer less ambiguity, which violates the independence axiom of the subjective expected utility model. Moreover, ambiguity tends to disappear when ambiguous events cannot be compared (Fox & Tversky, 1995).

Ambiguity has been widely investigated in the financial literature. Researchers have studied the effects of ambiguity on financial markets, portfolio choice, liquidity, and the 2008 financial crisis (Boyarchenko, 2012; Dow & Werlang, 1992; Routledge & Zin, 2009). A series of studies have also examined ambiguity in commodity futures markets, applying a theoretical approach (Lien & Wang, 2003; Lien & Yu, 2017; Wong, 2015).

The literature on ambiguity in financial and commodity futures markets is predominantly theoretical. The abstract nature of these theoretical discussions leaves a gap in an empirical approach; the challenge is to formulate a robust measure of the degree of ambiguity and ambiguity aversion for commodities markets. Commodity futures markets exhibit dynamics similar to financial markets. Thus, a measure of ambiguity and ambiguity aversion can help firms make optimal decisions regarding their production, hedging, and trading strategies.

We aim to estimate ambiguity in commodity futures markets by formulating an empirical ambiguity measure. We compose an empirical measure to explain 12 commodity prices of the nearest futures contracts. First, we calculate the maximum and minimum daily price spread average on the turnover by open interest to remove market noise, and we define the residuals as the measure of the degree of ambiguity. Second, we describe the interactions between the commodity log returns and ambiguity measure as a system, using an unrestricted vector autoregressive (VAR) framework to model the relationships between the variables. Finally, we categorize the results of the ambiguity impacts on the commodity futures markets' structural dynamics.

In sum, the contributions of the study are threefold. First, we formulate an empirical measure to evaluate ambiguity with real market data. Second, we investigate the effects of ambiguity on the dynamics of commodity futures markets. Third, we analyze the impact of ambiguity on futures price log returns. This is the first empirical research to develop ambiguity measures and analyze ambiguity impacts in commodity futures markets to the best of our knowledge.

5.2
Literature Review

5.2.1
Definition of uncertainty

The definition of uncertainty is illustrated in the Ellsberg paradox (Ellsberg, 1961). The Ellsberg paradox distinguishes between risk and "Knightian uncertainty," in which people tend to prefer known over unknown probabilities. In a hypothetical experiment, Ellsberg illustrates a decision between two urns containing red and black balls. Urn 1 has 100 red and black balls with an unknown ratio. Urn 2 contains 50 red and 50 black balls with a known ratio (i.e., the payoff is $100: if you bet on red and win, you get $100, and $0 otherwise). The likelihood of winning is $\frac{1}{2}\, n + (n-1)\frac{1}{2}$. However, the results from Ellsberg's experiment violate the independence of the subjective expected utility axiom, thereby showing an ambiguity aversion. Therefore, people prefer to bet on events when probabilities are known (Urn 2) rather than uncertain (Urn 1) (Camerer & Weber, 1992).

Previous studies show the implications of ambiguity for economic and financial behavior. Studies analyze the psychological causes and rationality of ambiguity (Frisch & Baron, 1988); the implications of ambiguity for portfolio choice puzzles (Dimmock et al., 2016); and applications of ambiguity in the design of bilateral economic contracts, the trade in financial contracts and financial markets, and strategic decision-making in auctions (Mukerji, 2000).

We define ambiguity as the uncertainty in the probability distribution. In the asset prices context, Dow et al.(1992) employ an expected utility under a nonadditive probability measure to discriminate risks and uncertainties based on the axiomatic foundation of the maxmin expected utility decision rule proposed by Gilboa and Schmeidler (1989). Agents consider the minimal expected utility over all priors in the set, and investors consider the worst-case scenario when lacking complete information about their probability distribution, reflecting the ambiguity aversion.

However, to discern between risk and ambiguity within a subjective expected utility framework and account for the "probabilistic sophistication," Chen and Epstein (2002) compose a continuous-time version of the multiple-prior utility with ambiguity aversion. The model can distinguish a premium for risk and a premium for ambiguity. Notably, Gilboa and Schmeidler's (1989) multiple-priors model has different implications for portfolio choice and asset price than subjective expected utility (Epstein & Schneider, 2010). Agents cannot discern between risky and ambiguous situations, but they can learn how to; thus, an increase in investors' confidence causes a trend toward stock market participation and investments (Epstein & Schneider, 2007). In developing a decision-maker preference, Klibanoff et al. (2005) separate the ambiguity, defined as a characteristic of the decision maker's subjective information, from risk. The authors identify that attitudes toward risk are characterized by the shape of a von Neumann–Morgenstern utility function.

Investors choose a set of conditional probabilities to maximize the expected utility in a worst-case scenario, and they react asymmetrically. Epstein and Schneider (2010) consider the role of uncertain information quality in financial markets when there is incomplete knowledge about signal quality; agents tend to react more to bad news than to good news. Thus, agents demand compensation for low future information quality when fundamentals are more volatile. However, estimating the quality of information is not a straightforward process. Ozsoylev et al. (2011) examine information transmission when the quality of the signals is unknown and consider how investors deal with ambiguous information in asset markets. Their investigation highlights the implications of market depth, liquidity risk, and trading volume. The authors elucidate that arbitrageurs provide market liquidity and do not trade when the information signal is ambiguous (i.e., in an illiquid market). Market depth is lower when ambiguity occurs, and the trading volume also decreases with ambiguity. Moreover, ambiguity can lead to price volatility; hence, illiquid markets can impact asset prices.

5.2.2
Measures of ambiguity

Investors can interpret and obtain information in various ways. Thus, the literature has developed different approaches to heterogeneous beliefs to capture different opportunity sets of information, subjective beliefs, and utility functions. Abel (1989) formulates a consumer model framework to examine the implications of heterogeneous subjective beliefs about asset prices and finds that an increase in heterogeneity can lead to a stock price reduction. Anderson et al. (2005) also analyze how heterogeneous beliefs impact asset prices and returns, using publicly available forecasts by financial analysts as a proxy for the heterogeneous agents' beliefs. Diether et al. (2002) show that stocks returns with higher dispersion in analysts' earnings forecasts—a proxy for divergent opinions among investors—are negatively related to futures returns, which contrasts with the assumption that dispersion in analysts' earnings forecasts is a proxy for risk. Regarding differences in investors' opinions, Hong et al. (2015) develop a theoretical model to explain market crashes. Earlier studies have used heterogeneous belief models to explain differences in investors' opinions about the financial market (Basak, 2000, 2005; Harris & Raviv, 1993; Wang, 1994; Zapatero, 1998).

Anderson et al. (2009) distinguish between risk and uncertainty; risky events have a known distribution with an unknown outcome, while uncertain events have both an unknown distribution and an unknown outcome. Hence, the authors use the level of agreement among professional forecasters to measure the degree of uncertainty in the mean return, the greater the agreement among forecasters, the lower the degree of uncertainty, and the greater the disagreement, the higher the uncertainty. The authors empirically define risk as volatility in an asset's return. Their findings indicate that risk is less important than uncertainty as a driver of expected market excess return; the degree of uncertainty correlates more closely with market excess return than risk.

Variability can be considered a proxy for uncertainty. For instance, Romer (1990) argues that the stock market crash in 1929 fueled an increase in uncertainty, leading consumers to delay their spending on durable goods. Bloom et al. (2007) define uncertainty as the standard deviation of daily stock returns; however, to avoid noise unrelated to fundamentals due to the volatility in stock returns, the authors normalize the firms' daily share returns on the Stock Market Index. Basu et al. (2017) and Bloom (2009) use the Chicago Board Options Exchange Volatility Index (VXO) as a measure of uncertainty. Similarly, Yang et al. (2000) gauge uncertainty based on the variance of an estimator: the smaller the variance, the more accurate the estimation. Their goal is to estimate the volatility by employing the opening, closing, high, and low prices for multiple periods.

Gauging uncertainty remains a challenge in the literature, but it can be approached using volatility as a proxy. Jurado et al. (2015) assert that the definition of uncertainty is the conditional volatility disturbance that cannot be predicted by economic agents. However, the ability to measure uncertainty depends on how the measure is correlated with the latent stochastic process. For instance, the authors explain that stock market volatility can change, while uncertainty about economic fundamentals remains steady. Thus, their work aims to distinguish between uncertainty and their own definition of conditional volatility.

Other studies have attempted to quantify ambiguity. Bossaerts et al (2010) employ an experiment based on observations that ambiguity has different implications for choices and prices. Tan et al. (2017) use a bid-ask spread to gauge ambiguity in the UK stock market. Izhakian (2020) measures ambiguity using the volatility of probabilities. The author states that the degree of ambiguity must be independent of the risk and the outcome. The relation between risk and return cannot be considered without ambiguity (Brenner & Izhakian, 2018). Using a static general equilibrium model, Augustin et al. (2019) investigate the relation of risk and ambiguity to credit default swaps (CDS) and show that ambiguity has a significant effect on spreads. Adopting a different approach, Baillon et al. (2018) develop an ambiguity measure of natural and artificial events in an experimental framework. Li (2017) applied (Baillon et al., 2018) work to study the relationship between people's ambiguity attitudes and income in a natural experiment.

Knightian uncertainty can also be applied to a hedging and production groundwork in which decisions can be made following a maxmin principle. Producers tend to engage in a one-to-one hedge ratio (Lien, 2000). For example, in a commodity futures market context, a producer would choose to hedge when the futures prices were higher than the spot price adjusted by ambiguity (Lien & Wang, 2003). In the presence of ambiguity, the opportunity to hedge is invaluable to the firm (Wong, 2015). Nevertheless, the timing of hedging plays a role in the production and hedging decisions by incorporating preferences into the competitive firm's behavior under price uncertainty (Lien & Yu, 2015). Hedging decisions can have different impacts on underpricing ambiguity when firms face optimistic versus pessimistic outcomes (Lien & Yu, 2017).

The commodities market experienced a profound transformation in the early 2000s. A substantial investment influx has entered the commodities market, resulting in a new asset class for many financial funds and reflecting a so-called financialization process (Tang & Xiong, 2012). Therefore, as a standard feature in the financial process, ambiguity should be investigated in the commodity market context.

5.3 Methods and data

5.3.1 Data

The dataset comprises high, low, and closing prices and open interest—the proxy of the futures nearest contract's daily volume. The initial date is January 3, 2000, and the end date is December 12, 2019, for a total of 5,203 observations. We examine 12 commodity futures prices: cocoa, coffee, corn, cotton, feeder cattle, lean hog, live cattle, orange juice, soybeans, sugar, wheat, and West Texas Intermediate (WTI) oil (www.bloomberg.com, 2020). We summarize the statistics for the daily ambiguity measure and the commodities price log returns in Appendix A.

Figure 9 illustrates the original data in log levels and the daily log returns for the 12 commodities futures markets, while Figure 10 depicts the log-level correlation between the involved variables. We notice the highest correlation between the feeder cattle and the live cattle and amid grains (i.e., corn and wheat).

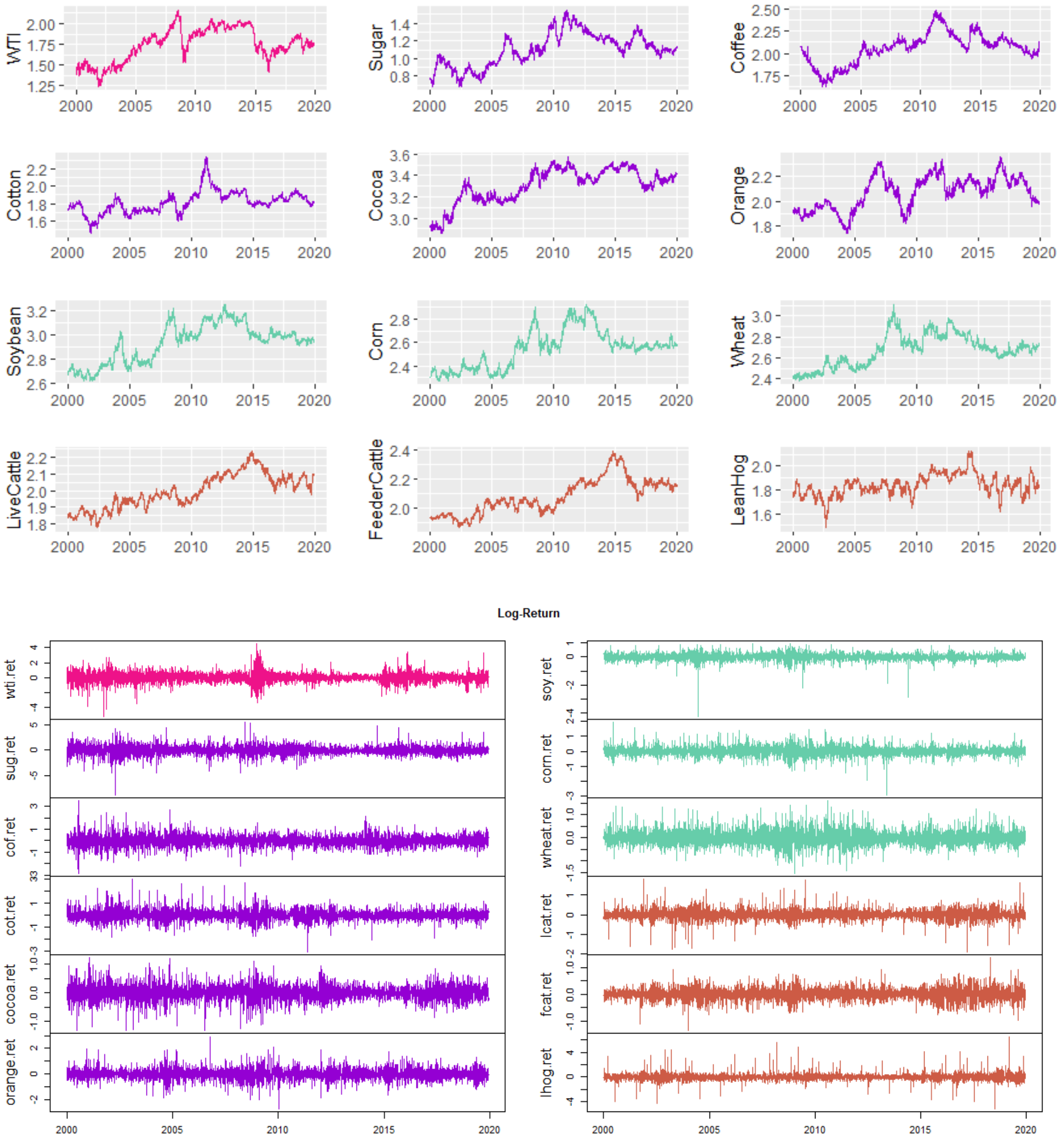


Figure 9. Commodity futures prices in levels and daily log returns. Period: January 3rd, 2000 to December 12th, 2019. 5.203 observations.

Source: Bloomberg.

Obs.: We split the 12 commodities by color to indicate the group: energy – WTI; softs - cocoa, coffee, cotton, and orange juice futures; grains and oilseeds – corn, soybeans, and wheat; livestock – feeder and live cattle and lean hog.

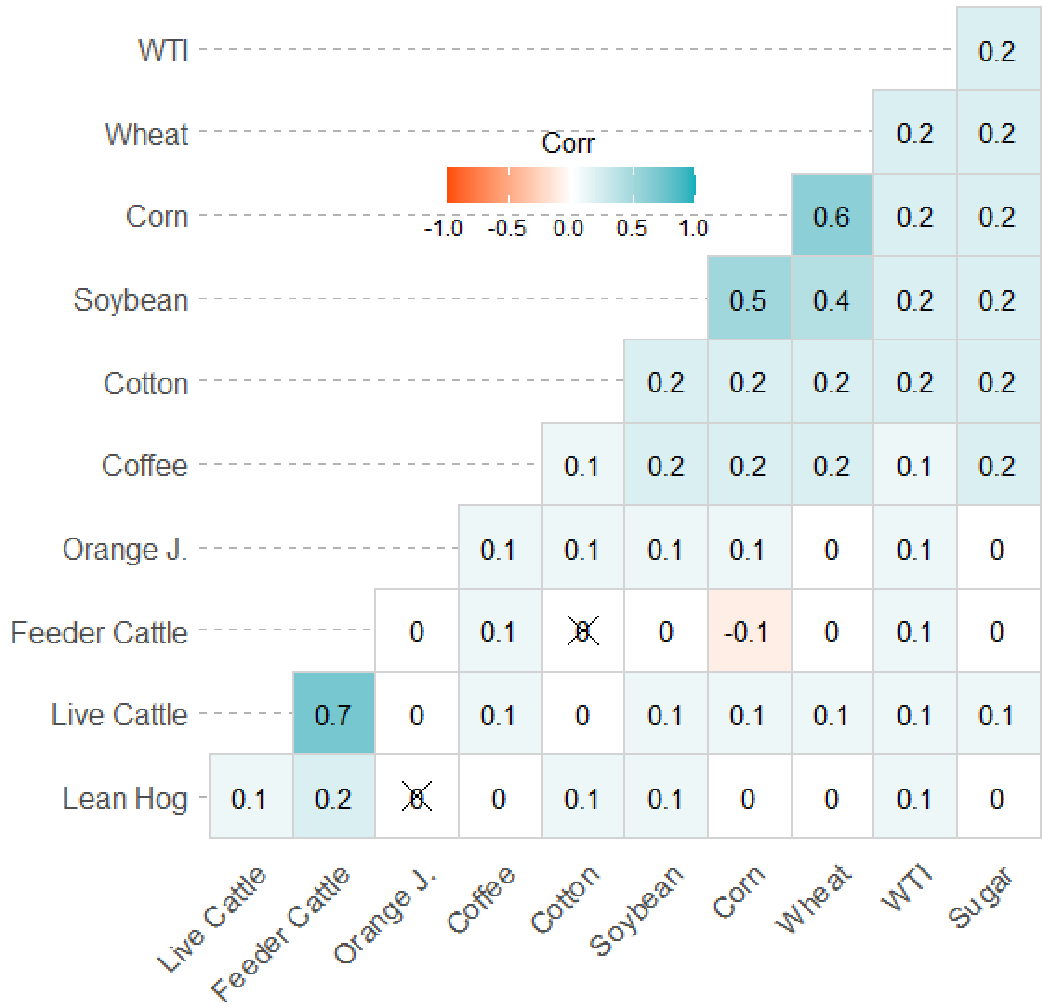


Figure 10. Correlations between agricultural commodities in log-levels.
Source: Research results.

Next, we test the UR using PP and ADF for the 12 commodities, ambiguity measure, and daily log returns in levels in Appendix B. For both the ambiguity measure and the log returns in levels, the PP and ADF UR tests reject the null hypothesis of one UR with a 1% statistical significance. The residual of the linear regression between the daily high-low mean price logarithms and open interest defines the ambiguity measure. The PP and ADF UR tests demonstrate that the ambiguity measure is a stationary covariance process. Thus, the PP and ADF UR tests of daily log returns are covariance stationary, in line with Baillie and Myers (1991).

All pairs of the return time series are I(0), so we can apply an unrestricted VAR framework to analyze the cross-effects of the time-series dynamics between the daily commodity prices ambiguity measure and log return without differencing the series.

5.3.2
Theoretical Framework

Gilboa and Schmeidler (1989) develop a preference relation over a set of deterministic outcomes (act): $J(f) = min\{\int u \circ fdP | P \in C\}$, where f is an act, u is a von Neumann–Morgenstern utility over outcomes, and C is a closed and convex set of finitely additive probability measures on the states of nature. The model allows a range of probabilities to be attributed to an outcome. The decision is based on a preference that is ranked according to the utility of the worst-case scenario. The decision process is generated by a minimization and maximization framework. Garlappi et al. (2007) found that the multi-priors model with ambiguity aversion achieves better results than the mean and variance approach.

We employ Gilboa and Schmeidler (1989) multiple-priors preference framework and Tan et al. (2017) approach, which first defines the degree of ambiguity in terms of the bid-ask spread. Since volatility indicates ambiguity in asset price dynamics (Epstein & Schneider, 2010), we formulate an empirical ambiguity measure. To estimate the bid-ask spread, we calculate a proxy using the high-low price spread. The advantage of the high-low price spread is the information about a financial asset's temporal dynamic process (Cheung & Chinn, 2001; Xiong, Li, & Bao, 2017). For example, the high-low price spread illustrates real traded prices, revealing the market makers' actions, the transaction costs, and the traders' degree of risk aversion.

To define an ambiguity measure, we assume that investors express homogenous preferences about the distribution of an asset return r^*. As a result of ambiguity, buyers and sellers estimate idiosyncratic worst-case scenarios and interpret the asset returns with different priors—the buyers' worst-case scenario is when the price falls after they take a long position; the opposite is true for sellers, whose worst-case scenario is when the price rises after they take a short position. Thus, buyers' and sellers' price discovery reveals the uncertainty in the information on futures prices, indicating ambiguity. As such, to bear the uncertainty of trade positions, buyers quote the lowest bid price whilst sellers quote the highest asking price, demanding a risk premium for this uncertainty. The risk premium expresses a possible tradeoff loss, holding the long and short positions if prices cross the asset return reference threshold, r^*.

If the degree of ambiguity of the returns is defined as k, $r^* - k$ expresses the buyers' prior return, and $r^* + k$ for the sellers. Specifically, to estimate the bid-ask daily spread, we use the daily high-low price spread proxy. We appraise the bid and ask prices B_t and A_t, as the low and high daily prices, L_t and H_t, defining the buyers' and sellers' ambiguity holding metrics, respectively, as:

$$ln\frac{L_t}{P_t} = lnL_t - lnP_t = r^* - k \qquad (1.1)$$

$$ln\frac{H_t}{P_t} = lnH_t - lnP_t = r^* + k \qquad (1.2)$$

where P_t expresses the commodity futures prices at time t. Subtracting equations, 1.2 and 1.1:

$$lnH_t - lnL_t = 2k \qquad (2)$$

Equation 2 defines a measure of the degree of ambiguity, k. Thus, we may apply Equation 3 as a proxy for the degree of ambiguity:

$$k = \frac{lnH_t - lnL_t}{2} \qquad (3)$$

We calculate the daily log returns of the commodity futures prices as:

$$r_t = ln\frac{P_t}{P_{t-1}} \qquad (4)$$

where r_t estimates the daily log return of the commodity futures prices at time t, and P_t expresses the commodity futures prices at time t.

We also need to assess the impact of market makers on the commodity futures price dynamics. Therefore, we regress the spread calculated using Equation 3 on a turnover by volume (we use open interest as a proxy for a daily volume of the futures contract) to remove the variation of the spread predicted by the market makers' impact.

The intricate task is to empirically separate the uncertainty price from the risk. To better capture the equation's uncertainty, it is crucial to understand the market makers' role, which is to offset price risk to customers and to facilitate risk-sharing. For instance, if the potential liquidity traders are net sellers, the market makers will charge a spread in the trade to offset the risk in operation, offering the sellers a price that is not "uncertain" (Grossman & Miller, 1988), when the number of market makers increases, the trade volume increases and the spread decreases (Biais et al., 2000). Furthermore, the wider the market markers' spread, the riskier and more uncertain the securities (Barnea & Logue, 1975). Tan et al.(2017) show that trading volume (m), number of market makers (n), the degree of risk-aversion (A), transaction cost (c), and volatility of the asset price σ^2 influence the liquidity (L) – which is indicated by the market makers' role in the financial market: $L = \frac{m}{n+1} + 2\frac{c.n}{A\sigma^2.(n+1)}$. They also show that the relationship between the trading volume and the bid-ask spread is not statistically significant in the ETF FTSE100 market, which can be explained by the high level of liquidity; hence, market makers do not explain the bid-ask spread. However, we expect to identify a relationship between the high-low price spread and the volume since commodity markets are not as liquid as the ETF market—agents would demand a risk premium for the uncertainty, and the spreads would open favoring the market markers' profit.

Another important aspect to consider is the volatility role in the formulation process. Epstein and Schneider (2010) highlight the importance of the interplay between volatility and ambiguity in asset pricing. Volatility is related to ambiguity in asset price dynamics, and it is also related to heterogeneous beliefs, but it is a consequence of those beliefs, not a proxy. Hibbert et al. (2020) shows empirical support in the literature, regardless of different sources of heterogeneous beliefs, higher investor disagreement leads to higher price uncertainty, and consequently, higher volatility return. In this sense, we preserve the volatility impact on the spread (Eq. 4) but remove the market maker effect on the high-low spread.

Furthermore, the residuals from a linear regression model can be applied to categorize the dependent and independent variables' relationships in a different context. For example, Bessembinder and Seguin (1993) analyze futures markets' expected returns, conditional on the week effect and volatility, by applying a linear regression. They use the residuals of the linear regression to identify unexpected returns. Bewley (2011) shows that each degree of ambiguity corresponds to a set of prior distributions over a Gaussian linear regression model's parameter. Chan and Fong (2000) employ the absolute values of return residuals in a two-stage regression to determine whether daily price volatility increases with trade size. Board et al. (2001) use the residuals of a linear bivariate traded spot and futures volume to measure information impact. The difference between the fitted values for the futures and spot markets gives a measure of information-less futures trading activity over information-less spot market activity.

We formulate Equation 5 as follows:

$$K_t = \alpha_0 + \widehat{\alpha_1} V_t + A_t \qquad (5)$$

where K_t = mean spread between H_t and L_t at time t; α_0 = constant; V_t = daily turnover volume at time t; and, A_t = residuals, a measure of the degree of ambiguity of the commodity futures prices, at time t.

We employ the volume to estimate the ambiguity measure; a trading volume may describe heterogeneous beliefs about market outcomes (Gallant, Rossi, & Tauchen, 1992; T. Li, 2007; Shalen, 1993). The literature indicates trading volume and ambiguity, which defines the lack of information about futures returns (Bewley, 2002; Dow & Werlang, 1992; Easley & O'Hara, 2010; Epstein & Schneider, 2007).

Bossaerts et al. (2010) analyze the ambiguity effects on market depth, liquidity risk, and trading volume and conclude that market depth is lower when agents identify ambiguity. Moreover, ambiguity decreases the trading volume decreases. Applying the same approach, we use the daily traded volume as the independent variable, and the daily high-low average spread as the dependent variable to calculate the linear regression residuals. Thus, we define the residuals as a proxy of the measure of ambiguity in commodity futures markets.

5.3.3
Preliminary assessment

To estimate the temporal dynamics, we use the residuals of an OLS regression between the degree of ambiguity, the average of the log of the high-low average spread, and open interest. However, the formulation may produce the two-step-estimation problem, resulting in consistent point estimators and inconsistent standard errors. The inconsistent standard errors problem can be overcome easily (Cragg, 1983).

To estimate robust standard errors in the OLS residuals between the average high-low spread and the volume, we apply the heteroskedasticity and autocorrelation (HAC) of unknown-form consistent covariances method by Newey and West (1987). In particular, the method compares the results of both the ordinary OLS and the HAC procedure standard errors. Additionally, the HAC robust Wald p-value is slightly higher than the corresponding non-robust F-statistic p-value, but both express statistical significance at conventional test levels. Noteworthy, the regression indicates that open interest (V_t) is statically significant to explain the high-low spread (K_t) for most of the commodities tested, except for WTI, Corn, and Wheat – which implies that market makers may contribute to the high-low spread in the less liquid markets (see Appendix C).

As a robustness test, we also estimate the unit root and the linear Granger causality between the OLS variables, revealing that the degree of ambiguity and the open interest are stationary. Consequently, the residuals express a linear combination of stationary variables, showing stationarity, as illustrated in the additional UR residual test. Therefore, the OLS does not describe spurious results.

Moreover, the Granger causality test identifies the existence of at least one-sided causal relationships between the dependent and independent OLS variables[6].

5.3.4
Vector Autoregression approach

We apply the unrestricted VAR framework to analyze the cross-effects between the degree of ambiguity and the daily log returns. First, we examine the covariance stationarity of the dependent variables using the Phillips–Perron (PP) and the augmented Dickey-Fuller (ADF) unit root (UR) tests for the ambiguity measure and daily log returns. If both series are stationary, I(0), we apply the unrestricted VAR model without differencing the series.

We formulate the VAR model for the commodity futures ambiguity measure and daily log returns as follows expressed in standard VAR form, which by hypothesis uses all variables as endogenous:

$$A_t = C_A + \hat{\alpha}_{A,t-1}r_{t-1} + \hat{\alpha}_{A,t-2}r_{t-2} + \cdots + \hat{\alpha}_{A,t-n}r_{t-n} + \hat{\beta}_{A,t-1}A_{t-1} + \hat{\beta}_{A,t-2}A_{t-2} + \cdots + \hat{\beta}_{A,t-n}A_{t-n} + \varepsilon_t^A$$

$$A_t = C_A + \textstyle\sum_{i=1}^{n} \hat{\alpha}_{t-i}r_{t-i} + \sum_{i=1}^{n} \hat{\beta}_{t-i}A_{t-i} + \varepsilon_t^A \qquad (6)$$

And,

$$r_t = C_r + \hat{\alpha}_{r,t-1}r_{t-1} + \hat{\alpha}_{r,t-2}r_{t-2} + \cdots + \hat{\alpha}_{r,t-n}r_{t-n} + \hat{\beta}_{r,t-1}A_{t-1} + \hat{\beta}_{r,t-2}A_{t-2} + \cdots + +\hat{\beta}_{r,t-n}A_{t-n} + \varepsilon_t^r$$

$$r_t = C_r + \textstyle\sum_{i=1}^{n} \hat{\alpha}_{t-i}r_{t-i} + \sum_{i=1}^{n} \hat{\beta}_{t-i}A_{t-i} + \varepsilon_t^r \qquad (7)$$

where r_t = daily log returns of commodity futures prices, at time t, defined by Equation 4; A_t = ambiguity measure of the commodity futures prices, at time t, represented by the residuals of Equation 5; n = number of VAR lags, estimated applying the Schwarz information criterion (SC); C_A and C_r = constants of the ambiguity measure and daily log returns, respectively; $\hat{\alpha}_{A,t-i}, \hat{\beta}_{A,t-i}$ and $\hat{\alpha}_{r,t-i}, \hat{\beta}_{r,t-i}$ express the VAR estimated coefficients for the daily ambiguity measure and the log returns, respectively; and, ε_t^A, ε_t^r = daily ambiguity measure and log return VAR errors.

Before applying Equations 6 and 7 for the commodity ambiguity measure and daily log return, we analyze the Johanssen cointegration trace and maximum eigenvalue tests to identify the number of cointegrating equations displayed in Appendix D. We analyze unrestricted VAR equations for A_t and r_t examining the individual coefficients' values, signals, and statistical significance.

Next, we generate the VAR impulse-response functions, illustrating pairwise the autoregressive (AR) and cross-effects between ambiguity and daily log returns. Each VAR produces the impulse response functions, which are two figures that describe the accumulated response to one standard deviation of Cholesky innovation of A_t and r_t, 10–14 days ahead of the shock period.

[6] For conciseness, the results are not reported here. However, they are available upon request.

Furthermore, we illustrate the pairwise linear Granger causality test between the commodity ambiguity measure and daily log returns. The Granger causality test solves bivariate regressions of the form for all possible pairs of (x, y) series in the group:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \cdots + \alpha_k y_{t-k} + \beta_1 x_{t-1} + \cdots + \beta_k x_{-k} + \epsilon_t \quad (8)$$

$$x_t = \alpha_0 + \alpha_1 x_{t-1} + \cdots + \alpha_k x_{t-k} + \beta_1 y_{t-1} + \cdots + \beta_k y_{-k} + u_t \quad (9)$$

The reported F-statistics are the Wald statistics for the joint hypothesis: $\beta_1 = \beta_2 = \ldots = \beta_k = 0$. For each equation, the null hypothesis is that x does not Granger-cause y in the first regression and that y does not Granger-cause x in the second regression. The Granger causality test shows the precedence and the information content but does not describe causality.

5.4
Results and Discussion

This section summarizes our research results. As previously outlined in the Methods and Data section, the ambiguity measure and the log return can reveal their dynamics, highlighting the different market depths and traders' profiles such as hedgers and speculators (i.e., noncommercial traders). The structural change caused by the increase in noncommercial traders may have decreased the risk premiums, cost of hedging, price volatility, and integrated commodity markets with financial markets (Irwin & Sanders, 2012b; Ozsoylev & Werner, 2011; Sanjuán-López & Dawson, 2017).

5.4.1
VAR framework

Different methods to estimate ambiguity have been proposed in the literature. Possible definitions of an ambiguity measure include volatility of total returns (Faria & Correia-da-Silva, 2014), volatility of mean returns (Franzoni, 2017), disagreement among analyst forecasts (Anderson et al., 2009), and expected volatility of probabilities (Izhakian, 2020). Given these diverse formulations, no empirical ambiguity measure can be considered universal.

To evaluate the ambiguity measure in commodity futures markets, we choose the multiple-priors preference mainstream approach adopted by Gilboa and Schmeidler (1989), Dow and Werlang (1992), and Garlappi et al. (2007), in combination with the model proposed by Tan et al. (2017). We define the ambiguity measure as the residuals of an OLS regression between the average high-low spread and the open interest. To analyze the ambiguity measure and log-return temporal dynamics in commodity markets, we apply an unrestricted vector autoregression (VAR) model.

The statistical significance value of the ambiguity measure at time t indicates the complete information set of the previous lags; thus, the ambiguity lags should not show statistical significance. However, our results show that the ambiguity measure lags express statistical significance for all commodities in all periods in the unrestricted VAR equation, indicating a robust autoregressive pattern. Depending on the commodity, the ambiguity measure lags show different degrees of statistical significance for the log-return VAR equation from zero to 1%, 5%, and 10% significance at multiple lags. Similar results are found in Tan et al. (2017): the VAR results for the log returns and ambiguity measure in the FTSE100 market demonstrate statistical significance. A possible explanation is the temporal persistence of the degree of ambiguity in the liquid FTSE100 market daily trading. Comparatively, for lower-liquidity commodity futures markets, we may distinguish a similar persistence pattern of the ambiguity degree resulting in the econometric estimation of statistically significant ambiguity lags.

In order to compose the contemporaneous and lagged endogenous variables relationship structure, we employ the unrestricted VAR framework. Specifically, we construct the equations for each pair of the 12 commodity prices ambiguity measure (A_t) and daily log return in levels (r_t), as shown in Table I.

Lags[1]	Cocoa A_t	Cocoa r_t	Coffee A_t	Coffee r_t	Corn A_t	Corn r_t	Cotton A_t	Cotton r_t	Feeder Cattle A_t	Feeder Cattle r_t	Lean Hog A_t	Lean Hog r_t
A_{t-1}	0.19***	0.02	0.17***	-0.05	0.24***	0.10**	0.28***	-0.99***	0.18***	0.01	0.19***	0.02
A_{t-2}	0.15***	0.02	0.12***	0.01	0.11***	-0.01	0.05***	0.42***	0.12***	0.04	0.09***	0.03
A_{t-3}	0.10***	-0.10**	0.08***	-0.02	0.08***	0.02	0.09***	0.05	0.06***	-0.03	0.08***	-0.01
A_{t-4}	0.09***	0.05	0.09***	0.02	0.08***	0.06	0.06***	0.03	0.10***	0.01	0.09***	0.07
A_{t-5}	0.06***	-0.01	0.09***	0.01	0.08***	0.01	0.05***	0.02	0.10***	-0.01	0.09***	0.07
A_{t-6}	0.09***	-0.03			0.08***	0.02	0.07***	0.09*	0.06***	0.03	0.08***	-0.06
A_{t-7}					0.08***	-0.04	0.06***	0.05	0.08***	0.02	0.05***	0.13*
A_{t-8}					0.08***	-0.14***	0.10***	0.06	0.09***	-0.01	0.07***	-0.12*
A_{t-9}											0.08***	-0.01
r_{t-1}	0.00	-0.00	0.03***	-0.03**	-0.00	0.03*	-0.00	0.22***	-0.04***	0.08***	-0.01***	-0.00
r_{t-2}	0.00	-0.01	0.03***	-0.03**	-0.00	-0.01	-0.00	-0.09***	-0.02***	-0.02	-0.01**	-0.00
r_{t-3}	0.00	0.02*	0.02***	0.02*	-0.00	-0.01	-0.01	0.00	-0.02***	0.00	-0.01***	-0.03**
r_{t-4}	0.00	-0.02	0.01**	-0.01	-0.00	-0.01	0.00**	0.01	-0.02***	0.01	-0.00	-0.03*
r_{t-5}	0.01*	0.01	0.01**	-0.01	-0.00	-0.01	-0.00	0.01	-0.00	-0.01	-0.00	0.01
r_{t-6}	0.00	-0.00			-0.00	-0.00	-0.00	-0.00	-0.00	-0.03**	-0.00	-0.01
r_{t-7}					-0.00	-0.02	-0.00	0.04***	-0.00	-0.00	-0.00	-0.00
r_{t-8}					-0.01	-0.00	-0.01*	-0.02	-0.00	-0.03*	-0.00	-0.00
r_{t-9}											0.01**	0.03*
c	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Obs.: 1. The numbers of lags defined by the Schwarz (SC) criterion. (*), (**) and (***) statistically significant at 1%, 5% and 10%, respectively.

Table I. Bi-variate VAR model. Commodity futures daily prices ambiguity measure (A_t) and log returns (r_t)

	Live Cattle		Orange Juice		Soybeans		Sugar		Wheat		WTI	
Lags[1]	A_t	r_t	A_t	r_t	A_t	r_t	A_t	r_t	A_t	r_t	A_t	r_t
A_{t-1}	0.17	-2.01	0.25***	-0.04	0.19***	-0.03	0.29***	-0.03	0.22*	-0.03	0.22***	-0.10**
A_{t-2}	1.50	3.01	0.11***	0.02	0.12***	-0.17***	0.12***	0.02	0.13*	-0.04	0.12***	0.07
A_{t-3}	-3.18	-5.95	0.11***	0.06	0.09***	0.08	0.10***	0.08*	0.08*	0.13***	0.10***	0.00
A_{t-4}	1.76	3.49	0.10***	0.08**	0.07***	-0.01	0.07***	0.07	0.09*	0.01	0.04***	-0.04
A_{t-5}	-0.72	-1.47	0.10***	-0.05	0.12***	-0.01	-0.14***	-0.07	0.07*	-0.08	0.12***	0.15***
A_{t-6}	2.10	4.26			0.08***	0.05			0.08*	0.02	0.00	0.04
A_{t-7}	-1.02	2.07			0.07***	0.09*			0.05*	-0.08*	0.00	0.07
A_{t-8}					0.08***	-0.07			0.08*	-0.03	0.03**	-0.03
A_{t-9}											-0.03*	-0.03
A_{t-10}											0.04	-0.11**
A_{t-11}											-0.02***	0.10*
A_{t-12}											0.04**	0.06
A_{t-13}											0.06***	-0.01
A_{t-14}											0.11***	0.10**
r_{t-1}	0.00	1.19	-0.00	0.07***	-0.00	-0.02	-0.01**	-0.01	0.03*	-0.00	-0.02***	-0.04***
r_{t-2}	-0.70	-1.40	0.00	-0.03**	-0.00	0.00	0.01**	-0.03***	0.02*	-0.01	-0.01*	-0.03**
r_{t-3}	1.63	3.05	-0.00	-0.03**	-0.00	-0.00	0.00	0.04*	0.01*	-0.02	-0.01***	-0.02*
r_{t-4}	-0.82	-1.62	0.00	-0.01	0.01*	0.01	0.00	-0.01	0.00	-0.01	-0.01	0.02
r_{t-5}	0.41	0.84	-0.01	0.00	0.00	-0.03*	0.00	0.01	0.00	0.00	-0.01**	-0.03**
r_{t-6}	-1.03	-2.07			0.01*	0.01			-0.00	0.00	-0.01***	-0.02**
r_{t-7}	0.55	1.12			-0.01	0.00			-0.00	-0.01	-0.01	-0.01
r_{t-8}					-0.01	0.00			-0.00	-0.01	-0.01	-0.01
r_{t-9}											-0.01***	-0.02
r_{t-10}											-0.01**	0.00
r_{t-11}											-0.02***	-0.02
r_{t-12}											-0.00	0.00
r_{t-13}											-0.00	0.03**
r_{t-14}											0.00	0.02
c	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Obs.: 1. The numbers of lags defined by the Schwarz (SC) criterion. (*), (**) and (***) statistically significant at 1%, 5% and 10%, respectively.

Table I. Bi-variate VAR model. Commodity futures daily prices ambiguity measure (A_t) and log returns (r_t) (Continuation).

For the sake of brevity, we divide the 12 commodities into four separate groups: (i) grains and oilseeds—corn, soybeans, and wheat; (ii) livestock—feeder and live cattle; (iii) softs—cocoa, coffee, cotton, and orange juice; and (iv) energy—WTI.

In the grain and oilseeds group, the ambiguity measure shows a significant AR pattern, while the log returns demonstrate no statistical significance or persistence. The price disagreements and ambiguous situations in the grains and oilseeds futures market indicate their dynamics, lasting 7–8 trading days. In the softs group, the price disagreements and ambiguous situations denote their dynamics, lasting 5–8 days. Thus, this lack of rich information regarding market depth can be associated with low trade volume, seasonality, and traders' profiles (commercial and noncommercial), displaying the ambiguity feature.

The livestock group results show a lack of rich information regarding ambiguity but not regarding the log returns' short-term dynamics. The price disagreements and ambiguous situations in the livestock futures markets indicate their dynamics, lasting 8–9 trading days.

WTI futures markets show a statistically AR pattern in both the VAR A_t and r_t equation coefficients. The unrestricted VAR expresses 14 statistically significant lags, the largest number of the 12 commodities analyzed.

Notably, ambiguity shows a significant AR pattern for all the commodity futures markets. Ambiguity-averse investors may demand an additional risk premium, illustrated by the lagged price structure.

In conclusion, the analysis of the unrestricted VAR equations for the 12 commodities shows that ambiguity expresses a current characteristic of the daily price series. Moreover, the degree of ambiguity—estimated by the ambiguity measure—demonstrates the varying degrees of AR patterns, persistence, and impact in the log returns identified by the coefficients' signals, values, and statistical significance. The log returns are also estimated to gauge the effect on the ambiguity measure and the AR pattern.

5.4.2
Granger causality test

We examine the statistical significance of the VAR lags combined with Table II, the Pairwise Granger Causality Test, between the ambiguity measures and log returns. In Table II, except for cocoa, we identify at least one Granger causality direction between the pair of variables, the ambiguity measure, and the log returns, explaining the forecasting robustness of one or both series from the lagged values. We also analyze the stability of VAR lags by plotting the AR characteristic polynomial's VAR inverse roots[7].

[7] The AR roots lie within the unit circle, indicating that all bivariate VARS lags show stability. Therefore, the unrestricted VAR models express stability, and the statistical significance of the lags illustrate the alternative of using the impulse response and the equations to forecast the current values of the ambiguity measure and the log returns. For conciseness, the results are not reported here. However, they are available upon request.

Commodity	Null Hypothesis (H_0)	
	$r_t \rightarrow A_t$	$A_t \rightarrow r_t$
Cocoa	0.88	1.15
Coffee	21.48^{***}	0.30
Corn	0.84	1.98^{**}
Cotton	1.38	68.73^{***}
Feeder cattle	18.81^{***}	0.28
Lean hog	5.72^{***}	1.14
Live cattle	0.43	1.45
Orange juice	0.65	2.02^{*}
Soybeans	2.01^{**}	2.66^{***}
Sugar	2.44^{**}	1.75
Wheat	10.21^{***}	1.91^{*}
WTI	6.30^{***}	2.37^{***}

Table II. Pairwise Granger Causality Tests.
Notes: Null Hypothesis (H0): Log return does not Granger cause ambiguity ($r_t \rightarrow A_t$). Ambiguity does not Granger cause log return ($A_t \rightarrow r_t$). (***) Statistically significant at 1%; (**) Statistically significant at 5%; and, (*) Statistically significant at 10%.

We distinguish different patterns of Granger causation for the ambiguity measure and log returns among the commodity groups. The main outliers are cocoa and live cattle, which show no statistically significant Granger directional causation between both series. Results can expose the speed of price adjustments in the market as well as the market depth differences among commodities markets.

The past values of the ambiguity measure indicate persistence. Notably, higher-order lags of ambiguity affect the returns, such as the results for the unrestricted VAR of corn, cotton, lean hog, soybeans, sugar, and WTI. However, an analysis of the rejection of the no-Granger causality hypothesis shows that either the ambiguity measure or the log return infers statistically significant forecasts about the variables in the unrestricted VAR model. So, if we accept the Granger causality hypothesis, the ambiguity measure can predict the commodity futures log return, and vice versa. The coefficients in the log returns equation suggest that past values may linearly forecast the daily returns of commodity futures markets, as well as volatility (Tan et al., 2017). Moreover, even if the ambiguity measure does not Granger-cause commodity futures returns, Epstein and Schneider (2010) demonstrate that ambiguity aversion may influence the risk premium of assets, resulting in higher volatility.

5.4.3
VAR impulse response

In Appendix C, we show the bivariate VAR impulse response graphs between the daily ambiguity measure and log returns for the 12 commodity futures markets, showing the accumulated response to one standard deviation of Cholesky innovation.

We categorize the results per group. For instance, in the softs group, we notice a mixed positive-to-null effect of the log return impulse on the ambiguity measure response and a positive impact of the ambiguity measure impulse on the log return response. The grains and oilseeds group illustrates similar intragroup effects for both the ambiguity measure and log returns. However, the livestock group shows an AR feature only for the ambiguity measure. Lastly, for the energy group, the WTI shows a positive and rising AR pattern for the ambiguity measure. In contrast, the WTI futures log returns impulse positively impacts the ambiguity measure response with a decreasing rate, and the ambiguity measure impulse shows a small effect on the log returns response.

The overall results expose the differences in production, storage, and commercialization cycles for the four groups. Notably, the ambiguity feature in commodities markets reveals the market agents' beliefs and priors regarding the underlying price distributions.

Commodity futures markets agents may use the ambiguity measure as a degree of liquidity. For instance, ambiguity-averse traders may update priors to identify potential trades. Additionally, ambiguity expresses impacts in some daily log returns and may be priced as an additional risk premium that affects portfolio selection or noncommercial traders' allocations. In this sense, ambiguity in commodity futures markets can be a useful tool to gauge market liquidity and to use as a strategic informational input.

The ambiguity measure provides strategic informational input for hedgers, speculators, traders, and producers. For example, the enhanced production and marketing strategies may benefit from lower cost-benefit information identified in the ambiguity measure, resulting in efficient economic resource allocation. Thus, reducing ambiguity has a potential beneficial impact on investors and firms (Easley & O'Hara, 2010).

5.5
Summary and Conclusions

We estimate ambiguity in 12 commodity futures markets and formulate an empirical ambiguity measure. First, we calculate a linear regression between the daily high-low average spread of the futures prices and the daily open interest to remove market noise, defining the residuals as the measure of the ambiguity degree. Second, we use an unrestricted VAR framework to examine the interactions between the commodity log returns and ambiguity measure as a system. Finally, we categorize the results of the ambiguity impacts on the structural dynamics of the commodity prices.

In particular, ambiguity demonstrates a current and statistically significant pattern of commodity futures markets. The ambiguity measure illustrates statistically significant AR behavior with varying degrees of standard deviation.

Possible explanations for ambiguity in commodity futures markets are the information richness, knowledge of the underlying price distributions, volume, market depth, storage, traders' profiles, and speed of futures market price adjustments expressed by the number of lags of each commodity futures unrestricted VAR equation of the ambiguity measure and log returns. Particularly, ambiguity may identify a consequence of the financialization of the commodity futures market, which attracts more trading volumes for storable commodities such as grain, oilseeds, softs, and oil in contrast to livestock.

As outlined in the literature, comprehension of ambiguity and the use of an ambiguity measure in commodity futures markets may enhance production, storage, trading, and hedging strategies. Specifically, ambiguity indicates a significant current feature of the dynamics of commodity futures prices, presenting an added and positive informational input for decision-makers and market participants.

Our study presents one limitation worth noting. Unrestricted VAR is unable to capture nonlinearities such as structural breaks and asymmetries in commodity futures prices (Andreou & Ghysels, 2002; Hamilton, 1989). Further research can investigate the cross-impact of ambiguity in the commodities and financial markets. We also suggest that future studies evaluate the ambiguity measure as a hedging strategy tool.

5.6 Appendix

	Cocoa		Coffee		Corn		Cotton	
	Ambiguity Measure	**Log Return**	**Ambiguity Measure**	**Log Return**	**Ambiguity Measure**	**Log Return**	**Ambiguity Measure**	**Log Return**
Mean	-5.31	0.00	-3.48	2.86	-3.29	0.00	9.09	5.13
Median	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00
Maximum	0.05	0.09	0.05	0.16	0.05	0.11	0.06	0.13
Minimum	-0.01	-0.10	-0.01	-0.13	-0.02	-0.19	-0.01	-0.20
Std. Dev.	0.01	0.02	0.01	0.02	0.00	0.02	0.01	0.02
Skewness	1.80	-0.19	1.67	0.22	1.65	-0.12	1.95	-0.05
Kurtosis	8.80	5.45	7.95	6.40	7.04	8.78	11.09	12.69
JB	10078	1327	7730	2551	5890	7247	17482	20326
Probability	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

	Feeder cattle		Lean hog		Live cattle		Orange juice	
	Ambiguity Measure	**Log Return**	**Ambiguity Measure**	**Log Return**	**Ambiguity Measure**	**Log Return**	**Ambiguity Measure**	**Log Return**
Mean	-1.43	9.77	-4.00	3.66	-1.27	0.01	7.88	3.00
Median	-0.00	0.00	-0.00	0.00	-0.00	0.01	-0.00	0.00
Maximum	0.03	0.07	0.05	0.28	0.02	0.06	0.05	0.15
Minimum	-0.00	-0.06	-0.01	-0.22	-0.01	0.00	-0.01	-0.14
Std. Dev.	0.00	0.01	0.01	0.02	0.00	0.01	0.01	0.02
Skewness	1.48	-0.12	1.59	1.12	1.35	1.34	1.70	0.05
Kurtosis	7.04	5.51	7.85	26.87	5.76	5.74	7.17	6.79
JB	5416	1378	7289	12464	3233	3184	6270	3115
Probability	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

	Soybeans		Sugar		Wheat		WTI	
	Ambiguity Measure	**Log Return**	**Ambiguity Measure**	**Log Return**	**Ambiguity Measure**	**Log Return**	**Ambiguity Measure**	**Log Return**
Mean	-2.10	0.00	7.79	0.00	1.71	0.00	-2.5	0.00
Median	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00
Maximum	0.04	0.06	0.06	0.14	0.10	0.11	0.10	0.16
Minimum	-0.02	-0.28	-0.01	-0.15	-0.01	-0.10	-0.02	-0.17
Std. Dev.	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02
Skewness	1.98	-1.98	1.93	-0.04	2.31	0.23	2.55	-0.11
Kurtosis	8.85	32.35	10.28	6.52	16.28	5.16	15.60	7.43
JB	10831	190194	14743	2695	42891	1057	40077	4260
Probability	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Source: Research results.

A. I. Descriptive statistics. Commodity futures prices ambiguity measure and daily log returns. Period: January 3rd, 2000 to December 12th, 2019. 5.203 observations.

Bloomberg Data	
Commodity	Ticker
WTI	CL1 Comdty
Sugar	SB1 Comdty
Soybean	S 1 Comdty
Corn	C 1 Comdty
Coffee	KC1 Comdty
Wheat	W 1 Comdty
Cotton	CT1 Comdty
Cocoa	CC1 Comdty
Live Cattle	LC1 Comdty
Feeder Cattle	FC1 Comdty
Lean Hog	LH1 Comdty
Orange J.	JO1 Comdty

A.II. Data – Bloomberg codes

Notes: the ticker corresponds to a Bloomberg code for each commodity

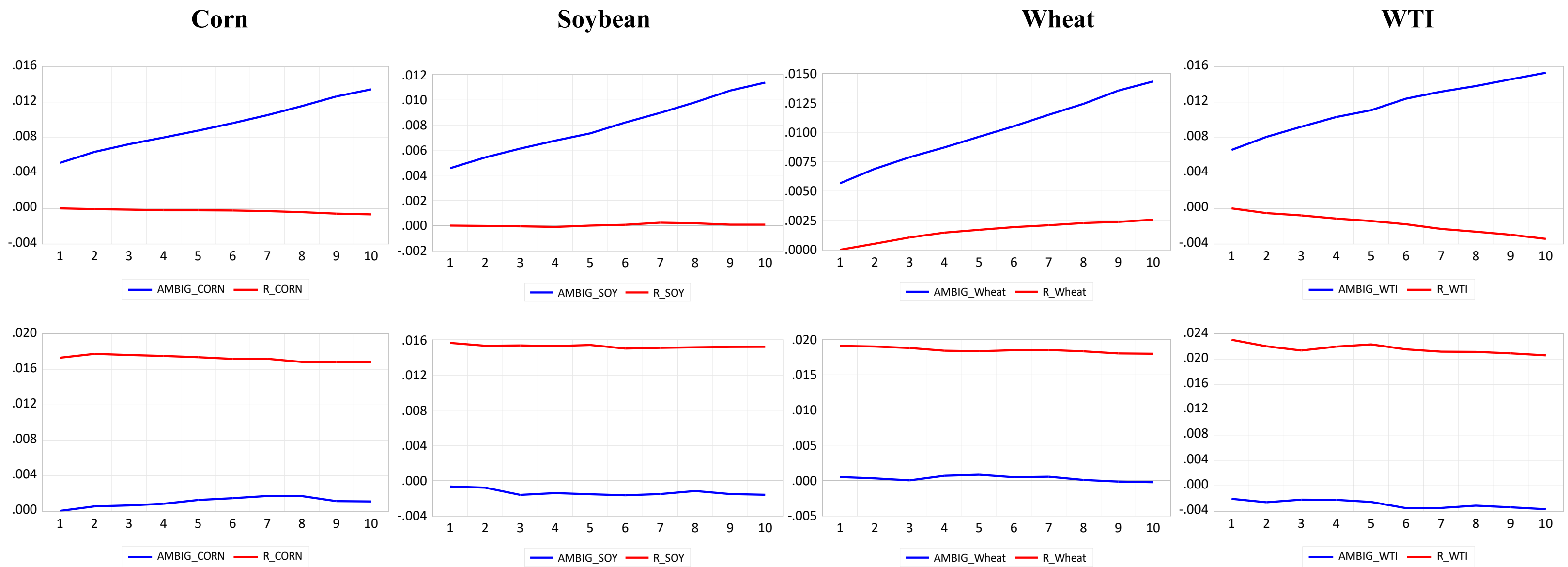
Unit root (UR) test				
Commodity	Ambiguity measure A_t		Log return r_t	
	PP	ADF	PP	ADF
Cocoa	-90.54***	-6.68***	-72.24***	-72.24***
Coffee	-79.32***	-15.17***	-74.60***	-74.29***
Corn	-83.42***	-10.29***	-70.19***	-70.21***
Cotton	-81.62***	-9.60***	-65.25***	-51.93***
Feeder cattle	-87.80***	-7.72***	-66.91***	-67.08***
Lean hog	-89.39***	-7.02***	-71.96***	-71.95***
Live cattle	-92.32***	-8.03***	-92.32***	-7.99***
Orange juice	-84.85***	-8.19***	-68.20***	-68.25***
Soybeans	-89.84***	-6.44***	-73.55***	-73.56***
Sugar	-82.11***	-5.28***	-72.98***	-41.08***
Wheat	-84.88***	-9.21***	-72.44***	-72.40***
WTI	-77.05***	-5.74***	-75.11***	-32.72***

B. I. Unit root (UR) tests. Phillips Perron (PP) and Augmented Dickey-Fuller (ADF). Commodity futures prices - ambiguity measure and daily log return in levels. Period: January 3rd, 2000 to December 12th, 2019. 5.203 observations (with constant).

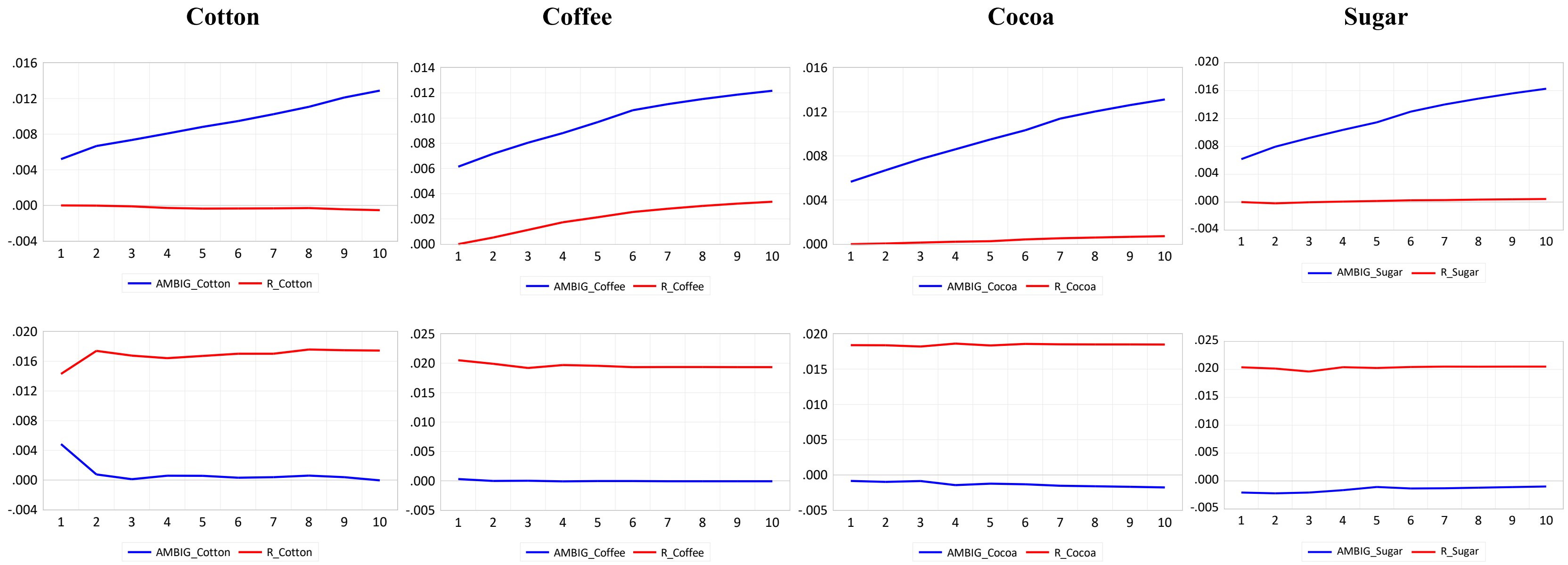
Source: Research results.

Obs.: (***) Statistically significant at 1%, rejecting the null hypothesis of unit root.

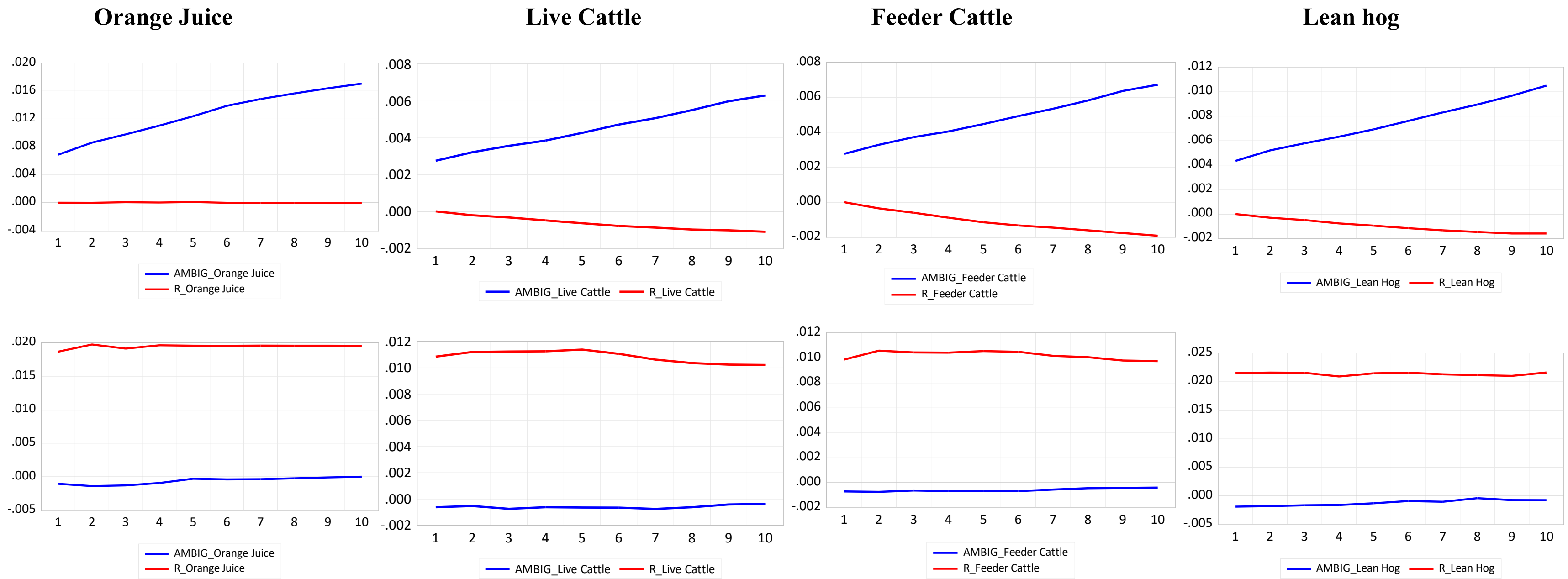
Appendix C.



C.I. Bi-variate VAR Impulse-response. Accumulated response to one SD of Cholesky innovation. Commodity futures daily ambiguity measures and log returns.
Period: January 3, 2000 to December 12, 2019. 5,203 observations.



C.II. Bi-variate VAR Impulse-response. Accumulated response to one SD of Cholesky innovation. Commodity futures daily ambiguity measures and log returns.



C.III. Bi-variate VAR Impulse-response. Accumulated response to one SD of Cholesky innovation. Commodity futures daily ambiguity measures and log returns.

6
Conclusion

This thesis sheds important light on the effects of the financialization of commodities markets on price discovery and how this process contributes to the price dynamics. We have seen the explosive behavior of the commodities prices in the 2007-2008 commodities boom and, more recently, the pandemic outbreak that pushed prices upward. Many investors resorted to the commodity market as an investment alternative.

There are four main conclusions that can be drawn from this thesis. First, we identified that speculation was not the main driver of the price return movements. Linear models showed that speculation tends to follow the prices' returns as well as the CIT long and short positions. However, nonlinear models were not conclusive, demonstrating the weakness in the relationship. Chapter one elucidated that one of the limitations of this model was that the weekly disclosure of the CFTC data report could mask the trading activities during the week. For instance, large investment funds may use different strategies and operate in the intraday position without apparent cause and effect. Moreover, larger traders who operate as commercial hedgers could take a speculative position with exclusive information.

The second chapter investigated the pass-through between international energy and feedstock commodities on the domestic price of gasoline and ethanol. The results demonstrated an asymmetric effect of prices for reformulated blendstock for oxygenate blending (RBOB) on gasoline prices in Brazil in both the short and long-term. The results also revealed that increases in heating oil (HO) prices lead to declines in the price of ethanol in the long run, and this supported the long-term conversion from fossil fuels to ethanol. This chapter highlighted the nexus between oil futures contracts in the international markets with the Brazilian fuel market. A wide range of stakeholders could benefit from these findings. Petrobras could adjust its price policy and its hedging strategy. Countries that depend on biofuel imports also could take advantage of these findings to lessen the effects of the spike in oil prices. Moreover, investors and traders could use this information to improve their hedging strategy.

Using the DY framework, the third chapter analyzed the volatility spillovers between the Brazilian agricultural commodity spot prices and futures markets. The DY framework illustrates ow to use a systematic application to examine volatility spillover, a useful tool for understanding the interconnectedness of international and local markets. This chapter showed the increasing connectedness over the year of futures and spot markets, intensifying after the pandemic outbreak. In addition, it has demonstrated the importance of the Brazilian soybean spot market after 2018, coincidentally after the China–US trade war.

The last chapter developed an ambiguity measure for commodity markets. The measurement demonstrated a current pattern in the commodity markets, showing that the degree of ambiguity changes depending on the market liquidity. Specifically, ambiguity indicates a significant current feature of the dynamics of commodity futures prices, presenting an added and positive informational input for decision-makers and market participants. The degree of ambiguity also highlights the consequence of the financialization process that attracts more trading volumes for storable commodities such as grain, oilseeds, softs and oil in contrast to livestock.

References

ABBOTT, P.; HURT, C.; TYNER, W. E. (2008). Whats Driving Food Prices? *Farm Foundation Issue Report*. https://doi.org/10.1016/j.gie.2008.05.046.

ABEL, A. B. (1989). Asset Prices under Heterogenous Beliefs: Implications for the Equity Premium. In *Working Paper, University of Pennsylvania* (Wharton School Rodney L. White Center for Financial Research).

AKAIKE, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705.

ALLEN, D. E.; CHANG, C.; MCALEER, M.; SINGH, A. K. (2018). A cointegration analysis of agricultural, energy and bio-fuel spot, and futures prices. *Applied Economics*, *50*(7), 804–823. https://doi.org/10.1080/00036846.2017.134058.

ALMEIDA, E. L. F. D. E.; OLIVEIRA, P. V. D. E.; LOSEKANN, L.; de Almeida, E. L. F.; de Oliveira, P. V.; LOSEKANN, L.; ALMEIDA, E. L. F. D. E.; OLIVEIRA, P. V. D. E.; LOSEKANN, L.; ALMEIDA, E. L. F. D. E.; OLIVEIRA, P. V. D. E.; LOSEKANN, L. (2015). Impactos da contenção dos preços de combustíveis no Brasil e opções de mecanismos de precificação. *Revista de Economia Politica*, *35*(3), 531–556. https://doi.org/10.1590/0101-31572015v35n03a09.

ANDERSON, E. W.; GHYSELS, E.; & JUERGENS, J. L. (2005). Do Heterogeneous Beliefs Matter for Asset Pricing? *The Review of Financial Studies*, *18*(3), 875–924.

ANDERSON, E. W.; GHYSELS, E.; & JUERGENS, J. L. (2009). The impact of risk and uncertainty on expected returns. *Journal of Financial Economics*, *94*(2), 233–263. https://doi.org/https://doi.org/10.1016/j.jfineco.2008.11.001.

ANDREASSON, P.; BEKIROS, S.; NGUYEN, D. K.; & UDDIN, G. S. (2016). Impact of speculation and economic uncertainty on commodity markets. *International Review of Financial Analysis*, *43*, 115–127. https://doi.org/10.1016/j.irfa.2015.11.005.

ANDREOU, E.; & GHYSELS, E. (2002). Detecting multiple breaks in financial market volatility dynamics. *Journal of Applied Econometrics*, *17*(5), 579–600. https://doi.org/10.1002/jae.684.

ANFAVEA (2020). *BRAZILIAN AUTOMOTIVE INDUSTRY YEARBOOK*. Brazilian Automotive Industry Association – ANFAVEA.

ANP (2019). *Síntese Mensal de Comercialização de Combustíveis*. http://www.anp.gov.br/arquivos/publicacoes/sinteses/2019-dezembro-sintese-volume.pdf.

ANP, Brazilian National Petroleum, N. G. and B. A. (2020). *Website*. http://www.anp.gov.br/

ANTONAKAKIS, N.; KIZYS, R. (2015). Dynamic spillovers between commodity and currency markets. *International Review of Financial Analysis*, *41*, 303–319. https://doi.org/https://doi.org/10.1016/j.irfa.2015.01.016.

ARAÚJO, W. A. (2016). *Chapter 1 - Ethanol Industry: Surpassing Uncertainties and Looking Forward* (S. L. M. Salles-Filho, L. A. B. Cortez, J. M. F. J. da Silveira, S. C. Trindade, & M. da G. D. B. T.-G. B. Fonseca, Eds.; pp. 1–33). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-803141-4.00001-0.

ATIL, A.; LAHIANI, A.; NGUYEN, D. K. (2014). Asymmetric and nonlinear pass-through of crude oil prices to gasoline and natural gas prices. *Energy Policy*, *65*, 567–573. https://doi.org/https://doi.org/10.1016/j.enpol.2013.09.064.

AUGUSTIN, P.; IZHAKIAN, Y. (2019). Ambiguity, Volatility, and Credit Risk. *The Review of Financial Studies*, *33*(4), 1618–1672. https://doi.org/10.1093/rfs/hhz082.

AULERICH, N. M.; IRWIN, S. H.; GARCIA, P. (2013). Bubbles, Food Prices, and Speculation: Evidence from the CFTC's Daily Large Trader Data Files. In *DeLong, J. B.; & Magin, K. (2006). A Short Note on the Size of the Dot-Com Bubble. National Bureau of Economic Research, (January), 1–14. http://doi.org/10.3386/w12011 NBER, J. H. and K. K. (2004). Lost Decade in Translation: Did the US Learn from Japan's* (Vol. w19065, Issue 58). https://doi.org/10.3386/w19065.

BABA, Y.; ENGLE, R.; KRAFT, D.; KRONER, K. (1990). *Multivariate simultaneous generalized arch, department of economics, university of california at san diego*. https://scholar.google.com.pk/scholar?hl=en&q=Baba%2C+Engle%2C+Kraft+and+Kroner+%281990%29+&btnG=&as_sdt=1%2C5&as_sdtp=.

BACHMEIER, L. J.; GRIFFIN, J. M. (2003). New Evidence on Asymmetric Gasoline Price Responses. *The Review of Economics and Statistics*, *85*(3), 772–776. http://www.jstor.org/stable/3211715.

BACON, R. W. (1991). Rockets and feathers: the asymmetric speed of adjustment of UK retail gasoline prices to cost changes. *Energy Economics*, *13*(3), 211–218. https://doi.org/10.1016/0140-9883(91)90022-R.

BAEK, E. G.; BROCK, W. A. (1992). *A general test for nonlinear Granger causality: Bivariate model*.

BAILLON, A.; HUANG, Z.; SELIM, A.; WAKKER, P. P. (2018). Measuring Ambiguity Attitudes for All (Natural) Events. *Econometrica*, *86*(5), 1839–1858. https://doi.org/10.3982/ECTA14370.

BALCILAR, M.; BEKUN, F. V. (2020). Do oil prices and exchange rates account for agricultural commodity market spillovers? Evidence from the Diebold and Yilmaz Index. *Agrekon*, *59*(3), 366–385. https://doi.org/10.1080/03031853.2019.1694046.

BALCOMBE, K.; RAPSOMANIKIS, G. (2008). Bayesian estimation and selection of nonlinear vector error correction models: The case of the sugar-ethanol-oil nexus in Brazil. *American Journal of Agricultural Economics*, *90*(3), 658–668. https://doi.org/10.1111/j.1467-8276.2008.01136.x.

BANERJEE, A.; DOLADO, J.; MESTRE, R. (1998). Error-correction Mechanism Tests for Cointegration in a Single-equation Framework. *Journal of Time Series Analysis*, *19*(3), 267–283. https://doi.org/10.1111/1467-9892.00091.

BARNEA, A.; LOGUE, D. E. (1975). The Effect of Risk on the Market Maker's Spread. *Financial Analysts Journal*, *31*(6), 45–49.

BASAK, S. (2000). A model of dynamic equilibrium asset pricing with heterogeneous beliefs and extraneous risk. *Journal of Economic Dynamics and Control*, *24*(1), 63–95. https://doi.org/https://doi.org/10.1016/S0165-1889(98)00064-5.

BASAK, S. (2005). Asset pricing with heterogeneous beliefs. *Journal of Banking & Finance*, *29*(11), 2849–2881. https://doi.org/https://doi.org/10.1016/j.jbankfin.2005.02.003.

BASU, S.; BUNDICK, B. (2017). Uncertainty Shocks in a Model of Effective Demand. *Econometrica*, *85*(3), 937–958. https://doi.org/10.3982/ECTA13960.

BEKIROS, S. (2014). Nonlinear causality testing with stepwise multivariate filtering: Evidence from stock and currency markets. *North American Journal of Economics and Finance*, *29*, 336–348. https://doi.org/10.1016/j.najef.2014.06.005.

BEKIROS, S. D.; DIKS, C. G. H. (2008a). The nonlinear dynamic relationship of exchange rates: Parametric and nonparametric causality testing. *Journal of Macroeconomics*, *30*(4), 1641–1650. https://doi.org/10.1016/j.jmacro.2008.04.001.

BEKIROS, S. D.; DIKS, C. G. H. (2008b). The relationship between crude oil spot and futures prices: Cointegration, linear and nonlinear causality. *Energy Economics*, *30*(5), 2673–2685. https://doi.org/10.1016/j.eneco.2008.03.006.

BESSEMBINDER, H.; SEGUIN, P. J. (1993). Price Volatility, Trading Volume, and Market Depth: Evidence from Futures Markets. *The Journal of Financial and Quantitative Analysis*, *28*(1), 21–39. https://doi.org/10.2307/2331149.

BEWLEY, T. F. (2002). Knightian decision theory. Part I. *Decisions in Economics and Finance*, *25*(2), 79–110. https://doi.org/10.1007/s102030200006.

BEWLEY, T. F. (2011). Knightian decision theory and econometric inferences. *Journal of Economic Theory*, *146*(3), 1134–1147. https://doi.org/https://doi.org/10.1016/j.jet.2010.12.012.

BIAIS, B.; MARTIMORT, D.; ROCHET, J.-C. (2000). Competing Mechanisms in a Common Value Environment. *Econometrica*, *68*(4), 799–837.

BLOOM, N. (2009). The Impact of Uncertainty Shocks. *Econometrica*, *77*(3), 623–685. https://doi.org/10.3982/ECTA6248.

BLOOM, N.; BOND, S.; VAN REENEN, J. (2007). Uncertainty and Investment Dynamics. *The Review of Economic Studies*, *74*(2), 391–415.

BOARD, J.; SANDMANN, G.; SUTCLIFFE, C. (2001). The Effect of Futures Market Volume on Spot Market Volatility. *Journal of Business Finance & Accounting*, *28*(7-8), 799–819. https://doi.org/10.1111/1468-5957.00394.

BOHL, M. T.; JAVED, F.; STEPHAN, P. M. (2012). Do Commodity Index Traders Destabilize Agricultural Futures Prices? *Ssrn*, *59*(2), 125–148. https://doi.org/10.2139/ssrn.2088426.

BOLLERSLEV, T.; ENGLE, R. F.; WOOLDRIDGE, J. M. (1988). A Capital Asset Pricing Model with Time-Varying Covariances. *Journal of Political Economy*, *96*(1), 116–131. https://doi.org/10.1086/261527.

BORENSTEIN, S.; CAMERON, A. C.; GILBERT, R. (1997). Do Gasoline Prices Respond Asymmetrically to Crude Oil Price Changes? *The Quarterly Journal of Economics*, *112*(1), 305–339.

BORENSTEIN, S.; SHEPARD, A. (2002). Sticky Prices, Inventories, and Market Power in Wholesale Gasoline Markets. *The RAND Journal of Economics*, *33*(1), 116–139. https://doi.org/10.2307/2696378.

BOSSAERTS, P.; GHIRARDATO, P.; GUARNASCHELLI, S.; ZAME, W. R. (2010). Ambiguity in Asset Markets: Theory and Experiment. *The Review of Financial Studies*, *23*(4), 1325–1359.

BOYARCHENKO, N. (2012). Ambiguity shifts and the 2007–2008 financial crisis. *Journal of Monetary Economics*, *59*(5), 493–507. https://doi.org/https://doi.org/10.1016/j.jmoneco.2012.04.002.

BRANCO, J. E. H.; BRANCO, D. H.; AGUIAR, E. M. DE, CAIXETA FILHO, J. V.; RODRIGUES, L. (2019). Study of optimal locations for new sugarcane mills in Brazil: Application of a MINLP network equilibrium model. *Biomass and Bioenergy*, *127*, 105249. https://doi.org/10.1016/j.biombioe.2019.05.018.

BRENNER, M.; IZHAKIAN, Y. (2018). Asset pricing and ambiguity: Empirical evidence∗. *Journal of Financial Economics*, *130*(3), 503–531. https://doi.org/https://doi.org/10.1016/j.jfineco.2018.07.007.

BROOCK, W. A.; SCHEINKMAN, J. A.; DECHERT, W. D.; LEBARON, B. (1996). A test for independence based on the correlation dimension. *Econometric Reviews*, *15*(3), 197–235. https://doi.org/10.1080/07474939608800353.

BRUNETTI, C.; BÜYÜKŞAHIN, B.; HARRIS, J. H. (2016). Speculators, Prices, and Market Volatility. *Journal of Financial and Quantitative Analysis*, *51*(5), 1545–1574. https://doi.org/10.1017/S0022109016000569.
BUMPASS, D.; GINN, V.; TUTTLE, M. H. (2015). Retail and wholesale gasoline price adjustments in response to oil price changes. *Energy Economics*, *52*, 49–54. https://doi.org/https://doi.org/10.1016/j.eneco.2015.08.030.

CAMERER, C.; WEBER, M. (1992). Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty*, *5*(4), 325–370. https://doi.org/10.1007/BF00122575.

CAPELLE-BLANCARD, G.; COULIBALY, D. (2011). Index Trading and Agricultural Commodity Prices: A Panel Granger Causality Analysis1 1The authors thank Agnès Bénassy-Quéré, Valérie Mignon and Sophie Piton for helpful comments. The usual disclaimer applies. *International Economics*, *126–127*, 51–71. https://doi.org/https://doi.org/10.1016/S2110-7017(13)60036-0.

CAVALCANTI, M.; SZKLO, A.; MACHADO, G. (2012). Do ethanol prices in Brazil follow Brent price and international gasoline price parity? *Renewable Energy*, *43*, 423–433. https://doi.org/https://doi.org/10.1016/j.renene.2011.11.034.

CEBALLOS, F.; HERNANDEZ, M. A.; MINOT, N.; ROBLES, M. (2017). Grain Price and Volatility Transmission from International to Domestic Markets in Developing Countries. *World Development*, *94*, 305–320. https://doi.org/https://doi.org/10.1016/j.worlddev.2017.01.015.

CEPEA (Center for Advanced Studies on Applied Economics). (2020). *Price Index*. http://www.cepea.esalq.usp.br/.

CFTC, C. F. T. C. (2008). *Commodity Swap Dealers & Index Traders with Commission Recommendations*.

CHAN, K.; FONG, W. M. (2000). Trade size, order imbalance, and the volatility–volume relation. *Journal of Financial Economics*, *57*(2), 247–273. https://doi.org/https://doi.org/10.1016/S0304-405X(00)00057-X.

CHENG, I.-H.; XIONG, W. (2014). Financialization of Commodity Markets. *Annual Review of Financial Economics*, *6*(1), 419–441. https://doi.org/10.1146/annurev-financial-110613-034432.

CHEUNG, Y. W.; CHINN, M. D. (2001). Currency traders and exchange rate dynamics: A survey of the US market. *Journal of International Money and Finance*, *20*(4), 439–471. https://doi.org/10.1016/S0261-5606(01)00002-X.

CHEVALLIER, J.; IELPO, F. (2013). Volatility spillovers in commodity markets. *Applied Economics Letters*, *20*(13), 1211–1227. https://doi.org/10.1080/13504851.2013.799748.

CHUA, C. L.; DE SILVA, C.; SUARDI, S. (2017). Do petrol prices increase faster than they fall in market disequilibria? *Energy Economics*, *61*, 135–146. https://doi.org/https://doi.org/10.1016/j.eneco.2016.10.024.
COSTA, C. C. DA, BURNQUIST, H. L. (2016). Impactos do controle do preÃ\So da gasolina sobre o etanol biocombustÃ\-vel no Brasil. *Estudos EconÃ\'{}micos (SÃ\poundso Paulo)*, *46*, 1003–1028.

CRAGG, J. G. (1983). More Efficient Estimation in the Presence of Heteroscedasticity of Unknown Form. *Econometrica*, *51*(3), 751. https://doi.org/10.2307/1912156.

CRAGO, C. L.; KHANNA, M.; BARTON, J.; GIULIANI, E.; AMARAL, W. (2010). Competitiveness of Brazilian sugarcane ethanol compared to US corn ethanol. *Energy Policy*, *38*(11), 7404–7415. https://doi.org/https://doi.org/10.1016/j.enpol.2010.08.016.

DAHL, R. E.; OGLEND, A.; YAHYA, M. (2020). Dynamics of volatility spillover in commodity markets: Linking crude oil to agriculture. *Journal of Commodity Markets*, *20*, 100111. https://doi.org/https://doi.org/10.1016/j.jcomm.2019.100111.

DAVID, S. A.; INÁCIO, C. M. C.; QUINTINO, D. D.; MACHADO, J. A. T. (2020). Measuring the Brazilian ethanol and gasoline market efficiency using DFA-Hurst and fractal dimension. *Energy Economics*, *85*, 104614. https://doi.org/https://doi.org/10.1016/j.eneco.2019.104614.

DAVID, S. A.; QUINTINO, D. D.; INACIO, C. M. C.; MACHADO, J. A. T. (2018). Fractional dynamic behavior in ethanol prices series. *Journal of Computational and Applied Mathematics*, *339*, 85–93. https://doi.org/https://doi.org/10.1016/j.cam.2018.01.007.

DE AZEVEDO, P. F.; SERIGATI, F. C. (2015). Preços administrados e discricionariedade do Executivo. *Revista de Economia Politica*, *35*(3), 510–530. https://doi.org/10.1590/0101-31572015v35n03a08.

DE GORTER, H.; DRABIK, D. (2016). Biofuel policies and the impact of developing countries' policy responses to the 2007–2008 food price boom. *Global Food Security*, *11*, 64–71. https://doi.org/https://doi.org/10.1016/j.gfs.2016.07.008.

DE OLIVEIRA, E. M.; CUNHA, F. A. F. S.; CYRINO OLIVEIRA, F. L.; SAMANEZ, C. P. (2017). Dynamic relationships between crude oil prices and socially responsible investing in Brazil: evidence for linear and non-linear causality. *Applied Economics*, *49*(22), 2125–2140. https://doi.org/10.1080/00036846.2016.1234695.

DE OLIVEIRA, E. M.; CYRINO OLIVEIRA, F. L.; KLÖTZLE, M. C.; PINTO, A. C. F. (2019). Dynamic Associations between GDP and Crude Oil Prices in Brazil: Structural Shifts and Nonlinear Causality. *Emerging Markets Finance and Trade*, 1–25. https://doi.org/10.1080/1540496X.2019.1658072.

DE OLIVEIRA, E. M.; DE SOUZA CUNHA, F. A. F.; PALAZZI, R. B.; KLOTZLE, M. C.; MAÇAIRA, P. M. (2020). On the effects of uncertainty measures on sustainability indices: An empirical investigation in a nonlinear framework. *International Review of Financial Analysis*, 101505. https://doi.org/https://doi.org/10.1016/j.irfa.2020.101505.

DENKER, M.; KELLER, G. (1983). On U-statistics and v. mise' statistics for weakly dependent processes. *Zeitschrift Für Wahrscheinlichkeitstheorie Und Verwandte Gebiete*, *64*(4), 505–522. https://doi.org/10.1007/BF00534953.

DICKEY, D. A.; FULLER, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association*, *74*(366), 427–431. https://doi.org/10.2307/2286348.

DIEBOLD, F. X.; LIU, L.; YILMAZ, K. (2017). Commodity Connectedness. In *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3038826.

DIEBOLD, F. X.; YILMAZ, K. (2009). Measuring Financial Asset Return and Volatility Spillovers, with Application to Global Equity Markets*. *The Economic Journal*, *119*(534), 158–171. https://doi.org/https://doi.org/10.1111/j.1468-0297.2008.02208.x.

DIEBOLD, F. X.; YILMAZ, K. (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*, *28*(1), 57–66. https://doi.org/https://doi.org/10.1016/j.ijforecast.2011.02.006.

DIEBOLD, F. X.; YILMAZ, K. (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, *182*(1), 119–134. https://doi.org/https://doi.org/10.1016/j.jeconom.2014.04.012.

DIETHER, K. B.; MALLOY, C. J.; SCHERBINA, A. (2002). Differences of Opinion and the Cross Section of Stock Returns. *The Journal of Finance*, *57*(5), 2113–2141. https://doi.org/10.1111/0022-1082.00490.

DIKS, C.; PANCHENKO, V. (2005). A note on the Hiemstra-Jones test for Granger non-causality. In *Studies in Nonlinear Dynamics and Econometrics* (Vol. 9, Issue 2). https://doi.org/10.2202/1558-3708.1234.

DIKS, C.; PANCHENKO, V. (2006). A new statistic and practical guidelines for nonparametric Granger causality testing. *Journal of Economic Dynamics and Control*, *30*(9–10), 1647–1669. https://doi.org/10.1016/j.jedc.2005.08.008.

DIMMOCK, S. G.; KOUWENBERG, R.; MITCHELL, O. S.; PEIJNENBURG, K. (2016). Ambiguity aversion and household portfolio choice puzzles: Empirical evidence. *Journal of Financial Economics*, *119*(3), 559–577. https://doi.org/https://doi.org/10.1016/j.jfineco.2016.01.003.

DOW, J.; WERLANG, S. R. DA C. (1992). Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolio. *Econometrica*, *60*(1), 197–204. https://doi.org/10.2307/2951685.

DRABIK, D.; DE GORTER, H.; JUST, D. R.; TIMILSINA, G. R. (2014). The Economics of Brazil's Ethanol-Sugar Markets, Mandates, and Tax Exemptions. *American Journal of Agricultural Economics*, *97*(5), 1433–1450. https://doi.org/10.1093/ajae/aau109.
EASLEY, D.; O'HARA, M. (2010). Microstructure and ambiguity. *Journal of Finance*, *65*(5), 1817–1846. https://doi.org/10.1111/j.1540-6261.2010.01595.x.

ELLIOTT, G.; ROTHENBERG, T. J.; STOCK, J. H. (1996). Efficient Tests for an Autoregressive Unit Root. *Econometrica*, *64*(4), 813–836. https://doi.org/10.2307/2171846.

ELLSBERG, D. (1961). Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics*, *75*(4), 643–669. https://doi.org/10.2307/1884324.

ENGLE, R. F.; GRANGER, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, *55*(2), 251–276. https://doi.org/10.2307/1913236.

ENGLE, R. F.; KRONER, K. F. (1993). MULTIVARIATE SIMULTANEOUS GENERALIZED ARCH Multivariate Simultaneous Generalized ARCH*. In *Econometric Theory* (Vol. 11, Issue 1). Cambridge University Press. http://www.jstor.org/stable/3532933.

ENGLE, R. F.; KRONER, K. F. (1995). Multivariate Simultaneous Generalized Arch. *Econometric Theory*, *11*(1), 122–150.

EPSTEIN, L. G.; SCHNEIDER, M. (2007). Learning under Ambiguity. *The Review of Economic Studies*, *74*(4), 1275–1303.

EPSTEIN, L. G.; SCHNEIDER, M. (2010). Ambiguity and Asset Markets. *Annual Review of Financial Economics*, *2*(1), 315–346. https://doi.org/10.1146/annurev-financial-120209-133940.

ETIENNE, X. L.; IRWIN, S. H.; GARCIA, P. (2018). Speculation and corn prices. *Applied Economics*, *50*(44), 4724–4744. https://doi.org/10.1080/00036846.2018.1466992.

FARIA, G.; CORREIA-DA-SILVA, J. (2014). A closed-form solution for options with ambiguity about stochastic volatility. *Review of Derivatives Research*, *17*(2), 125–159. https://doi.org/10.1007/s11147-014-9097-9.

FARKAS, R.; YONTCHEVA, B. (2019). Price transmission in the presence of a vertically integrated dominant firm: Evidence from the gasoline market. *Energy Policy*, *126*, 223–237. https://doi.org/https://doi.org/10.1016/j.enpol.2018.11.016.

FASANYA, I. O.; ODUDU, T. F. (2020). Modeling return and volatility spillovers among food prices in Nigeria. *Journal of Agriculture and Food Research*, *2*, 100029. https://doi.org/https://doi.org/10.1016/j.jafr.2020.100029.

FERNANDEZ-PEREZ, A.; FRIJNS, B.; TOURANI-RAD, A. (2016). Contemporaneous interactions among fuel, biofuel and agricultural commodities. *Energy Economics*, *58*, 1–10. https://doi.org/https://doi.org/10.1016/j.eneco.2016.05.014.

FIGUEROLA-FERRETTI, I.; GILBERT, C. L.; MCCRORIE, J. R. (2015). Testing for Mild Explosivity and Bubbles in LME Non-Ferrous Metals Prices. *Journal of Time Series Analysis*, *36*(5), 763–782. https://doi.org/https://doi.org/10.1111/jtsa.12121.

FIGUEROLA-FERRETTI, I.; MCCRORIE, J. R.; PARASKEVOPOULOS, I. (2020). Mild explosivity in recent crude oil prices. *Energy Economics*, *87*, 104387. https://doi.org/https://doi.org/10.1016/j.eneco.2019.05.002.

FOX, C. R.; TVERSKY, A. (1995). Ambiguity Aversion and Comparative Ignorance*. *The Quarterly Journal of Economics*, *110*(3), 585–603. https://doi.org/10.2307/2946693.

FRISCH, D.; BARON, J. (1988). Ambiguity and rationality. *Journal of Behavioral Decision Making*, *1*(3), 149–157. https://doi.org/10.1002/bdm.3960010303
Furtado, A. T.; Scandiffio, M. I. G.; & Cortez, L. A. B. (2011). The Brazilian sugarcane innovation system. *Energy Policy*, *39*(1), 156–166. https://doi.org/https://doi.org/10.1016/j.enpol.2010.09.023.

GALEOTTI, M.; LANZA, A.; MANERA, M. (2003). Rockets and feathers revisited: an international comparison on European gasoline markets. *Energy Economics*, *25*(2), 175–190. https://doi.org/https://doi.org/10.1016/S0140-9883(02)00102-0.

GALLANT, A. R.; ROSSI, P. E.; TAUCHEN, G. (1992). Stock Prices and Volume. *The Review of Financial Studies*, *5*(2), 199–242.

GARLAPPI, L.; UPPAL, R.; WANG, T. (2007). Portfolio selection with parameter and model uncertainty: A multi-prior approach. *Review of Financial Studies*, *20*(1), 41–81. https://doi.org/10.1093/rfs/hhl003.

GILBERT, C. L. (2010). How to Understand High Food Prices. *Journal of Agricultural Economics*, *61*(2), 398–425. https://doi.org/10.1111/j.1477-9552.2010.00248.x.

GILBERT, C. L.; MUGERA, H. K. (2020). Competitive Storage, Biofuels and the Corn Price. *Journal of Agricultural Economics*, *71*(2), 384–411. https://doi.org/https://doi.org/10.1111/1477-9552.12362.

GILBERT, C. L.; PFUDERER, S. (2014). *The financialization of food commodity markets* (Issue 6, pp. 122-148 BT-Handbook on Food). Edward Elgar Publishing. https://econpapers.repec.org/RePEc:elg:eechap:14879_6.

GILBOA, I.; SCHMEIDLER, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, *18*(2), 141–153. https://doi.org/https://doi.org/10.1016/0304-4068(89)90018-9.

GRANGER, C. W. J. (1980). Testing for causality. A personal viewpoint. *Journal of Economic Dynamics and Control*, *2*(C), 329–352. https://doi.org/10.1016/0165-1889(80)90069-X.

GRANGER, C. W. J. (1988). Causality, cointegration, and control. *Journal of Economic Dynamics and Control*, *12*(2–3), 551–559. https://doi.org/10.1016/0165-1889(88)90055-3.

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

GROSSMAN, S. J.; MILLER, M. H. (1988). Liquidity and Market Structure. *The Journal of Finance*, *43*(3), 617–633. https://doi.org/10.2307/2328186.

GUILLEMINOT, B.; OHANA, J. J.; OHANA, S. (2014). The interaction of speculators and index investors in agricultural derivatives markets. In *Agricultural Economics (United Kingdom)* (Vol. 45, Issue 6). https://doi.org/10.1111/agec.12122.

GUTIERREZ, L. (2013). Speculative bubbles in agricultural commodity markets. *European Review of Agricultural Economics*, *40*(2), 217–238. https://doi.org/10.1093/erae/jbs017
Hallack, L. N.; Kaufmann, R.; & Szklo, A. S. (2020). Price discovery in Brazil: causal relations among prices for crude oil, ethanol, and gasoline. *Energy Sources, Part B: Economics, Planning, and Policy*, *15*(4), 230–251. https://doi.org/10.1080/15567249.2020.1780346.

HAMILTON, J. D. (1989). A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica*, *57*(2), 357–384. https://doi.org/10.2307/1912559.

HAMMOUDEH, S.; LAHIANI, A.; NGUYEN, D. K.; SOUSA, R. M. (2015). An empirical analysis of energy cost pass-through to CO2 emission prices. *Energy Economics*. https://doi.org/10.1016/j.eneco.2015.02.013.

HAMMOUDEH, S.; LI, H.; JEON, B. (2003). Causality and volatility spillovers among petroleum prices of WTI, gasoline and heating oil in different locations. *The North American Journal of Economics and Finance*, *14*(1), 89–114. https://doi.org/https://doi.org/10.1016/S1062-9408(02)00112-2.

HARRIS, M.; RAVIV, A. (1993). Differences of Opinion Make a Horse Race. *The Review of Financial Studies*, *6*(3), 473–506.

HIBBERT, A. M.; KANG, Q.; KUMAR, A.; MISHRA, S. (2020). Heterogeneous beliefs and return volatility around seasoned equity offerings. *Journal of Financial Economics*, *137*(2), 571–589. https://doi.org/https://doi.org/10.1016/j.jfineco.2020.03.003.

HIEMSTRA, C.; JONES, J. D. (1994). Testing for Linear and Nonlinear Granger Causality in the Stock Price- Volume Relation. *The Journal of Finance*, *49*(5), 1639–1664. https://doi.org/10.2307/2329266.

HONARVAR, A. (2009). Asymmetry in retail gasoline and crude oil price movements in the United States: An application of hidden cointegration technique. *Energy Economics*, *31*(3), 395–402. https://doi.org/https://doi.org/10.1016/j.eneco.2009.01.010.

HONG, H.; & STEIN, J. C. (2015). Differences of Opinion, Short-Sales Constraints, and Market Crashes. *The Review of Financial Studies*, *16*(2), 487–525. https://doi.org/10.1093/rfs/hhg006.

HSIEH, D. A. (1991). Chaos and Nonlinear Dynamics: Application to Financial Markets. *The Journal of Finance*, *46*(5), 1839–1877. https://doi.org/10.1111/j.1540-6261.1991.tb04646.x.

IRWIN, S. H. (2013). Commodity index investment and food prices: Does the “masters hypothesis” explain recent price spikes? *Agricultural Economics (United Kingdom)*, *44*(SUPPL1), 29–41. https://doi.org/10.1111/agec.12048.

IRWIN, S. H.; SANDERS, D. R. (2011). Index Funds, Financialization, and Commodity Futures Markets. *Applied Economic Perspectives and Policy*, *33*(1), 1–31. http://www.jstor.org.proxy2.library.illinois.edu/stable/41237206.

IRWIN, S. H.; SANDERS, D. R. (2012a). Testing the Masters Hypothesis in commodity futures markets. *Energy Economics*, *34*(1), 256–269. https://doi.org/10.1016/j.eneco.2011.10.008.

IRWIN, S. H.; SANDERS, D. R. (2012b). Financialization and Structural Change in Commodity Futures Markets. *Journal of Agricultural and Applied Economics*, *44*(3), 371–396. https://doi.org/DOI: 10.1017/S1074070800000481.

IRWIN, S. H.; SANDERS, D. R.; MERRIN, R. P. (2016). Devil or Angel? The Role of Speculation in the Recent Commodity Price Boom (and Bust). *Journal of Agricultural and Applied Economics*, *41*(02), 377–391. https://doi.org/10.1017/s1074070800002856

Izhakian, Y. (2020). A theoretical foundation of ambiguity measurement. *Journal of Economic Theory*, *187*, 105001. https://doi.org/https://doi.org/10.1016/j.jet.2020.105001.

JÚNIOR, G. S. R. G. D. S. R.; PALAZZI, R. B. R. B.; KLOTZLE, M. C. M. C.; PINTO, A. C. F. A. C. F. (2019). Analyzing herding behavior in commodities markets – an empirical approach. *Finance Research Letters*, *65*, 84–93. https://doi.org/10.1016/j.frl.2019.08.033.

JURADO, K.; LUDVIGSON, S. C.; NG, S. (2015). Measuring uncertainty. *American Economic Review*, *105*(3), 1177–1216. https://doi.org/10.1257/aer.20131193
Kapetanios, G.; Shin, Y.; & Snell, A. (2003). Testing for a unit root in the nonlinear STAR framework. *Journal of Econometrics*, *112*(2), 359–379. https://doi.org/https://doi.org/10.1016/S0304-4076(02)00202-6.

KISSWANI, K. M. (2019). Asymmetric gasoline-oil price nexus: recent evidence from non-linear cointegration investigation. *Applied Economics Letters*, *26*(21), 1802–1806. https://doi.org/10.1080/13504851.2019.1602701.
KLIBANOFF, P.; MARINACCI, M.; MUKERJI, S. (2005). A Smooth Model of Decision Making under Ambiguity. *Econometrica*, *73*(6), 1849–1892. https://doi.org/10.1111/j.1468-0262.2005.00640.x.

KOOP, G.; PESARAN, M. H.; POTTER, S. M. (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics*, *74*(1), 119–147. https://doi.org/https://doi.org/10.1016/0304-4076(95)01753-4.

KRISTOUFEK, L.; JANDA, K.; ZILBERMAN, D. (2014). Price transmission between biofuels, fuels, and food commodities. *Biofuels, Bioproducts and Biorefining*, *8*(3), 362–373. https://doi.org/10.1002/bbb.1464.

KWIATKOWSKI, D.; PHILLIPS, P. C. B.; SCHMIDT, P.; SHIN, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, *54*(1), 159–178. https://doi.org/https://doi.org/10.1016/0304-4076(92)90104-Y.

LEHECKA, G. V. (2015). Do hedging and speculative pressures drive commodity prices, or the other way round? *Empirical Economics*, *49*(2), 575–603. https://doi.org/10.1007/s00181-014-0886-7.

LEYBOURNE, S. Y.; LLOYD, T. A.; REED, G. V. (1994). The excess comovement of commodity prices revisited. *World Development*, *22*(11), 1747–1758. https://doi.org/https://doi.org/10.1016/0305-750X(94)00081-6.

LI, C. (2017). Are the poor worse at dealing with ambiguity? *Journal of Risk and Uncertainty*, *54*(3), 239–268. https://doi.org/10.1007/s11166-017-9262-2.

LI, T. (2007). Heterogeneous beliefs, asset prices, and volatility in a pure exchange economy. *Journal of Economic Dynamics and Control*, *31*(5), 1697–1727. https://doi.org/https://doi.org/10.1016/j.jedc.2006.06.003.

LIEN, D. (2000). Production and hedging under Knightian uncertainty. *Journal of Futures Markets*, *20*(4), 397–404. https://doi.org/10.1002/(SICI)1096-9934(200004)20:4<397::AID-FUT6>3.0.CO;2-J.

LIEN, D.; WANG, Y. (2003). Futures market equilibrium under Knightian uncertainty. *Journal of Futures Markets*, *23*(7), 701–718. https://doi.org/10.1002/fut.10081.

LIEN, D.; YU, C.-F. (Jeffrey). (2015). Production and Anticipatory Hedging under Time-Inconsistent Preferences. *Journal of Futures Markets*, *35*(10), 961–985. https://doi.org/10.1002/fut.21697.

LIEN, D.; YU, C.-F. (Jeffrey). (2017). Production and hedging with optimism and pessimism under ambiguity. *International Review of Economics & Finance*, *50*, 122–135. https://doi.org/https://doi.org/10.1016/j.iref.2017.04.001.

MAITAH, M.; PROCHÁZKA, P.; SMUTKA, L.; MAITAH, K.; HONIG, V. (2019). Analysis of the Impact of Ethanol Production on Agricultural Product Prices in Brazil. *Sugar Tech*, *21*(5), 773–779. https://doi.org/10.1007/s12355-019-00709-w.
MCKAY, B.; SAUER, S.; RICHARDSON, B.; HERRE, R. (2016). The political economy of sugarcane flexing: initial insights from Brazil, Southern Africa and Cambodia. *The Journal of Peasant Studies*, *43*(1), 195–223. https://doi.org/10.1080/03066150.2014.992016.

MELO, A. S.; SAMPAIO, Y. S. B. (2014). Impactos dos preços da gasolina e do etanol sobre a demanda de etanol no Brasil. *Revista de Economia Contemporanea*, *18*(1), 56–83. https://doi.org/10.1590/141598481813.

MITCHELL, D. (2008). *A note on rising food prices: Vol. World Bank*. https://doi.org/10.1596/1813-9450-4682.

MUKERJI, S. (2000). A survey of some applications of the idea of ambiguity aversion in economics. *International Journal of Approximate Reasoning*, *24*(2), 221–234. https://doi.org/https://doi.org/10.1016/S0888-613X(00)00036-0.

NASCIMENTO FILHO, A. S.; SABA, H.; DOS SANTOS, R. G. O.; CALMON, J. G. A.; ARAÚJO, M. L. V.; JORGE, E. M. F.; MURARI, T. B. (2021). Analysis of Hydrous Ethanol Price Competitiveness after the Implementation of the Fossil Fuel Import Price Parity Policy in Brazil. *Sustainability*, *13*(17), 9899. https://doi.org/10.3390/su13179899.

NAZLIOGLU, S. (2011). World oil and agricultural commodity prices: Evidence from nonlinear causality. *Energy Policy*, *39*(5), 2935–2943. https://doi.org/10.1016/j.enpol.2011.03.001.

NAZLIOGLU, S.; ERDEM, C.; SOYTAS, U. (2013). Volatility spillover between oil and agricultural commodity markets. *Energy Economics*, *36*, 658–665. https://doi.org/https://doi.org/10.1016/j.eneco.2012.11.009.

OECD, FOOD; OF THE UNITED NATIONS, A. O. (2019a). *OECD-FAO Agricultural Outlook 2019-2028*. https://doi.org/https://doi.org/https://doi.org/10.1787/agr_outlook-2019-en.

OECD, FOOD; OF THE UNITED NATIONS, A. O. (2019b). *OECD-FAO Agricultural Outlook (Edition 2019)*. https://doi.org/https://doi.org/https://doi.org/10.1787/eed409b4-en.

OLADUNJOYE, O. (2008). Market structure and price adjustment in the U.S. wholesale gasoline markets. *Energy Economics*, *30*(3), 937–961. https://doi.org/https://doi.org/10.1016/j.eneco.2007.09.009.

OLIVEIRA, S. M., RIBEIRO, C. O.; CICOGNA, M. P. V. (2018). Uncertainty effects on production mix and on hedging decisions: The case of Brazilian ethanol and sugar. *Energy Economics*, *70*, 516–524. https://doi.org/10.1016/j.eneco.2018.01.025.

OZSOYLEV, H.; & WERNER, J. (2011). Liquidity and asset prices in rational expectations equilibrium with ambiguous information. *Economic Theory*, *48*(2), 469. https://doi.org/10.1007/s00199-011-0648-0.

PALAZZI, R. B.; FIGUEIREDO PINTO, A. C.; KLOTZLE, M. C. M. C.; OLIVEIRA, E. M. (2020). Can we still blame index funds for the price movements in the agricultural commodities market? *International Review of Economics & Finance*, *65*, 84–93. https://doi.org/https://doi.org/10.1016/j.iref.2019.10.001.

PESARAN, H. H.; SHIN, Y. (1998). Generalized impulse response analysis in linear multivariate models. *Economics Letters*, *58*(1), 17–29. https://doi.org/https://doi.org/10.1016/S0165-1765(97)00214-0.

PESARAN, H.; SMITH, R.; SHIN, Y. (2001). Bound Testing Approaches to the Analysis of Level Relationship. *Journal of Applied Econometrics*, *16*, 289–326. https://doi.org/10.1002/jae.616.

PHILLIPS, P. C. B.; PERRON, P. (1988). Testing for a Unit Root in Time Series Regression. *Biometrika*, *75*(2), 335–346. https://doi.org/10.2307/2336182.

PINDYCK, R. S.; ROTEMBERG, J. J. (1990). The Excess Co-Movement of Commodity Prices. *The Economic Journal*, *100*(403), 1173. https://doi.org/10.2307/2233966.

PIOTROWSKI, J.; COADY, D.; TYSON, J.; OSSOWSKI, R.; GILLINGHAM, R.; TAREQ, S. (2010). Petroleum Product Subsidies: Costly, Inequitable, and On the Rise. *IMF Staff Position Notes*, *2010*(05), 1. https://doi.org/10.5089/9781455239108.004.

POLEMIS, M. L.; TSIONAS, M. G. (2016). An alternative semiparametric approach to the modelling of asymmetric gasoline price adjustment. *Energy Economics*, *56*, 384–388. https://doi.org/https://doi.org/10.1016/j.eneco.2016.04.004.

QUINTINO, D. D.; DAVID, S. A.; & VIAN, C. E. de F. (2017). Analysis of the Relationship between Ethanol Spot and Futures Prices in Brazil. In *International Journal of Financial Studies* (Vol. 5, Issue 2). https://doi.org/10.3390/ijfs5020011.

QUINTINO, D. D.; FERREIRA, P. (2021). Diesel prices in Brazil: A dynamic fractional integration analysis. *Economics and Business Letters*, *10*(2 SE-), 116–125. https://doi.org/10.17811/ebl.10.2.2021.116-125.

RADCHENKO, S. (2005). Oil price volatility and the asymmetric response of gasoline prices to oil price increases and decreases. *Energy Economics*, *27*(5), 708–730. https://doi.org/https://doi.org/10.1016/j.eneco.2005.06.001.

RFA. (2020). *Renewable Fuels Association*. Statistics. http://www.ethanolrfa.org/pages/statistics.

ROBLES, M.; TORERO, M.; BRAUN, J. VON. (2009). When speculation matters: International Food Policy Research Institute, Issue Brief 57. In *IFPRI - Issue Brief* (Issue No.57). International Food Policy Research Institute. https://www.cabdirect.org/cabdirect/abstract/20093129784.

ROMER, C. D. (1990). The Great Crash and the Onset of the Great Depression. *The Quarterly Journal of Economics*, *105*(3), 597–624. https://doi.org/10.2307/2937892
ROUTLEDGE, B. R.; ZIN, S. E. (2009). Model uncertainty and liquidity. *Review of Economic Dynamics*, *12*(4), 543–566. https://doi.org/https://doi.org/10.1016/j.red.2008.10.002.

SALLES-FILHO, S. L. M.; CASTRO, P. F. D. DE, BIN, A.; EDQUIST, C.; FERRO, A. F. P.; CORDER, S. (2017). Perspectives for the Brazilian bioethanol sector: The innovation driver. *Energy Policy*, *108*, 70–77. https://doi.org/10.1016/j.enpol.2017.05.037.

SANDERS, D. R.; IRWIN, S. H. (2017). Bubbles, Froth and Facts: Another Look at the Masters Hypothesis in Commodity Futures Markets. *Journal of Agricultural Economics*, *68*(2), 345–365. https://doi.org/10.1111/1477-9552.12191.

SANDERS, D. R.; IRWIN, S. H.; MERRIN, R. P. (2010). The Adequacy of Speculation in Agricultural Futures Markets: Too Much of a Good Thing? *Applied Economic Perspectives and Policy*, *32*(1), 77–94.

SANDERSRS, D. R.; IRWIN, S. H. (2011). The impact of index funds in commodity futures markets: A systems approach. *Journal of Alternative Investments*, *14*(1), 40–49. https://doi.org/10.3905/jai.2011.14.1.040.

SANJUÁN-LÓPEZ, A. I.; DAWSON, P. J. (2017). Volatility Effects of Index Trading and Spillovers on US Agricultural Futures Markets: A Multivariate GARCH Approach. *Journal of Agricultural Economics*, *68*(3), 822–838. https://doi.org/10.1111/1477-9552.12216.

SARWAR, M. N.; HUSSAIN, H.; MAQBOOL, M. B. (2020). Pass through effects of oil price on food and non-food prices in Pakistan: A nonlinear ARDL approach. *Resources Policy*, *69*, 101876. https://doi.org/https://doi.org/10.1016/j.resourpol.2020.101876.

SCHWARZ, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461–464.

SERRA, T.; ZILBERMAN, D.; GIL, J. (2010). Price volatility in ethanol markets. *European Review of Agricultural Economics*, *38*(2), 259–280. https://doi.org/10.1093/erae/jbq046.

SHALEN, C. T. (1993). Volume, Volatility, and the Dispersion of Beliefs. *The Review of Financial Studies*, *6*(2), 405–434. https://doi.org/10.1093/rfs/6.2.405.

SIMS, C.; STOCK, J.; W WATSON, M. (1990). Inference in Linear Time Series Models with Some Unit Roots. In *Econometrica* (Vol. 58). https://doi.org/10.2307/2938337.

TAN, R. Q.; MANAHOV, V.; THIJSSEN, J. (2017). The effect of ambiguity on the UK stock market: Evidence from a new empirical approach. *Investment Management and Financial Innovations*, *14*(4), 133–147. https://doi.org/10.21511/imfi.14(4).2017.12.

TANG, K.; XIONG, W. (2012). Index investment and the financialization of commodities. *Financial Analysts Journal*, *68*(6), 54–74. https://doi.org/10.2469/faj.v68.n6.5.
TIMMER, C. P. (2008). Causes of high food prices. In *ADB Economics Working Paper Series* (Vol. 128).

TODA, H. Y.; YAMAMOTO, T. (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics*, *66*(1–2), 225–250. https://doi.org/10.1016/0304-4076(94)01616-8.

TROSTLE, R. (2008). *Global Agricultural Supply and Demand : Factors contributing to recent increases in food commodity prices Food commodity prices rose 70 % since January 2006*.

VACHA, L.; JANDA, K.; KRISTOUFEK, L.; ZILBERMAN, D. (2013). Time–frequency dynamics of biofuel–fuel–food system. *Energy Economics*, *40*, 233–241. https://doi.org/https://doi.org/10.1016/j.eneco.2013.06.015.

VERCAMMEN, J.; DOROUDIAN, A. (2014). Portfolio Speculation and Commodity Price Volatility in a Stochastic Storage Model. *American Journal of Agricultural Economics*, *96*(2), 517–532. https://doi.org/10.1093/ajae/aat098.

WANG, J. (1994). A Model of Competitive Stock Trading Volume. *Journal of Political Economy*, *102*(1), 127–168.

WONG, K. P. (2015). Ambiguity and the Value of Hedging. *Journal of Futures Markets*, *35*(9), 839–848. https://doi.org/10.1002/fut.21678.

WORKING, H. (1960). Speculation on Hedging Markets. *Food Research Institute Studies*, *Volume 01*(Number 2,), 136578. https://econpapers.repec.org/article/agsfrisst/136578.htm.

WRIGHT, B. (2014). Global biofuels: Key to the puzzle of grain market behavior. In *Journal of Economic Perspectives* (Vol. 28, Issue 1). https://doi.org/10.1257/jep.28.1.73.

WRIGHT, B. D. (2011). The economics of grain price volatility. *Applied Economic Perspectives and Policy*, *33*(1), 32–58. https://doi.org/10.1093/aepp/ppq033.

XIONG, T.; LI, C.; BAO, Y. (2017). Interval-valued time series forecasting using a novel hybrid HoltI and MSVR model. *Economic Modelling*, *60*, 11–23. https://doi.org/10.1016/j.econmod.2016.08.019.

YANG, D.; ZHANG, Q. (2000). Drift-independent volatility estimation based on high, low, open, and close prices. *Journal of Business*, *73*(3), 477–491. https://doi.org/10.1086/209650.

YOON, S.-M.; AL MAMUN, M.; UDDIN, G. S.; KANG, S. H. (2019). Network connectedness and net spillover between financial and commodity markets. *The North American Journal of Economics and Finance*, *48*, 801–818. https://doi.org/https://doi.org/10.1016/j.najef.2018.08.012.

ZAPATERO, F. (1998). Effects of financial innovations on market volatility when beliefs are heterogeneous. *Journal of Economic Dynamics and Control*, *22*(4), 597–626. https://doi.org/https://doi.org/10.1016/S0165-1889(97)00076-6.

ZHANG, D.; BROADSTOCK, D. C. (2020). Global financial crisis and rising connectedness in the international commodity markets. *International Review of Financial Analysis*, *68*, 101239. https://doi.org/https://doi.org/10.1016/j.irfa.2018.08.003.