Understanding and Forecasting the Effects of Global Shocks on Fuel Prices: The Brazilian case*

Bruno Tebaldi Q. Barbosa† Pedro Valls ‡ Emerson Marçal §

4 de abril de 2025

Resumo

Global shocks in oil markets are transmitted to the final price of gasoline throughout many channel markets. Investigating and measuring these mechanisms is an important issue. The Brazilian case is analyzed because a large part of the country has a car fleet equipped with flexible engines that can be fueled with alcohol or gasoline - 4 out 5 cars have this technology - and there is a wide supply of gasoline and alcohol all over the country to the consumer in the filling station. This work uses GVAR Global VAR methodology extended by impulse indicator saturation techniques to analyze the relationship between fuel prices between alcohol and gasoline. The methodology used in the paper allows us to identify outliers and evaluate how changes in the oil price affect gasoline and alcohol prices in different regions, and their lags and allows gains in terms of forecast accuracy.

Keywords: GVAR, cointegration, breaks, IIS, regional dynamics, Oil, Fuel market.

*The authors acknowledge the partial financial support given by São Paulo Research Foundation (FAPESP) in the grant number 2023/01728-0. The opinions, hypotheses, conclusions, or recommendations expressed in this material are solely the responsibility of the authors and may not necessarily reflect the views of FAPESP.

† B^3 and EESP-FGV. E-mail: bruno.barbosa@fgv.br

‡ Director of Center of Quantitative Studies in Economics and Finance E-mail: pedro.valls@fgv.br

§ Head of CEMAP -EESP-FGV, Lecturer at EESP-FGV. E-mail: emerson.marcal@fgv.br

1 Introduction

Gasoline accounts for about half the U.S. consumption of petroleum products, and its price is the most visible among these products (Balke et al., 1998). As such, changes in gasoline prices are constantly being studied and observed very carefully. Following the outbreak of the Gulf crisis on 2 August 1990, crude oil prices rose dramatically, resulting in the fourth price 'shock' since 1973; this brought to attention the response of retail gasoline prices to fluctuations in world oil prices (Shin, 1994). In theory, price shocks can originate at any point from crude oil prices to the final price at the gasoline pump. Shocks created at the wholesale price of gasoline may reflect a bottleneck in distribution. In contrast, price shocks originating farther upstream are more likely to represent the effects of variation in crude oil supply. Given the history of oil supply shocks and indications that gasoline demand is relatively stable, intuition suggests that price shocks are more likely to originate upstream and be transmitted downstream.

The effects of a shock to crude oil in retail gasoline depend on the response in many intermediate channels. Extensive literature has argued an asymmetric relationship between gasoline and oil prices, specifically that gasoline prices respond more quickly when oil prices are rising than when oil prices are falling. Moreover, some of this literature has found heterogeneous asymmetry in gas price responses across cities (Bennett et al., 2021).

Regardless of the symmetry of the relationship, it is clear that macroeconomic shocks affect sectors and regions with different intensities and delays. Sector activities tend to be concentrated in some areas or cities. The existence of good economic infrastructures such as roads, airports, ports, and telecommunications facilities links regions economically and provides channels that can spread out economic shocks among areas. Macroeconomic models tend to disregard regional and spatial economic activity issues due to their aggregate nature bias or complexity of theme. However, for a specific region, it is essential to understand how economic activity variance can be decomposed into macroeconomic and idiosyncratic factors.

In the past few years, time-series data at disaggregated and regional levels and long-period samples become available. At the same time, efforts have been made in econometric literature to tackle high-dimensional problems. These challenges must be faced to address interdependence among regions and linkages from the macroeconomic environment to regions. The first is the curse of dimensionality. The second challenge involves modelling the connection between regions and regional shocks.

One can tackle the curse of dimensionality in different ways, reducing the parameter space or shrinkage of the number of variables. The parameter space can be shrunk by imposing a set of restrictions, which could be, for instance, obtained from a theoretical structural model, directly on the parameters (e.g., Lasso model proposed by Tibshirani (1996); Adaptive Lasso Model presented by Zou (2006)). For the data shrinkage, one could use techniques where prior distributions are imposed on the parameters to be estimated. Bayesian Vector Autoregression model proposed by Doan et al. (1984), for example, use what has become known as 'Minnesota' priors to shrink the parameters space. The global vector autoregressive (GVAR) model proposed by Pesaran et al. (2004), is a model that can effectively tackle the interdependence among regions while handling the curse of dimensionality by shrinkage of the data.

Since the mid-1970s, there has been a growing interest in utilizing alcohol as an alternative and more environmentally friendly fuel for internal combustion engines. Recently, the use of alcohol as an alternative fuel has gained significance due to its minimal adverse effects on the environment. Unlike various fossil fuels and their exhaust emissions, which contribute to harmful environmental impacts like carbon monoxide, carbon dioxide, hydrocarbons, nitrogen oxides, and particulate matter, alcohol-based fuels offer a cleaner and greener alternative (Kasibhatta, 2019).

Brazil has a long tradition on using alcohol as an alternative to gasoline as fuel for cars. The initial effort to foster the use of alcohol was taken in 1975, as a response to the shock caused by the first oil crisis. The Brazilian government implemented the National Alcohol Program called "Pró-Álcool"[1], a nationwide program financed by the government to phase out automotive fuels derived from fossil fuels in favor of ethanol made from sugar cane. After extensive research in the 90s, the Brazilian subsidiary of Volkswagen launched the first fully flexible-fuel car in March of 2006. By 2010, most Brazilian automakers were producing models of flex cars and light trucks. The adoption of ethanol flex-fuel vehicles was so successful that the production of flex cars went from almost 40 thousand in 2003 to 1.7 million in 2007 (ANFAVEA, 2010).

During the first decade of 2000 cars manufactured in Brazil started to be produced with flexible engines that could use either alcohol or gasoline. The decision of choosing alcohol or gasoline was transferred from car acquisition to the filling stations and the reactions to shocks in the price of alcohol and gasoline were faster and stronger. The customer can direct arbitrage between the two fuels.

Therefore, this study proposes using the GVAR methodology to model the fuel prices at a regional level, accounting for interdependence among regions and global macroeconomic variables such as crude oil and exchange rate. The model GVAR model used in this study is referred to the GVAR-IIS, which is the classical global vector autoregressive extended with the inclusion of impulse-saturation dummies. The GVAR methodology eliminates the curse of dimensionality. It provides a parsimonious model for the fuel prices. The dummies saturation allows the model to consider any structural break or regional shocks, which can originate at any point from crude oil prices to the final price at the fuel pump.

In our empirical analysis, we focus on modelling data from Brazil, a vast country

[1]Portuguese: Programa Nacional do Álcool

with diverse regional dynamics. The Brazilian economy has been subject to significant macroeconomic shocks during our analysis period, making it an intriguing case for model evaluation. This study focuses its attention on the gasoline and ethanol market within Brazil, with three primary objectives. First and foremost, we aim to model the dynamics of regional fuel markets, considering the interplay between different Brazilian regions, regional shocks, and macroeconomic variables. To accomplish this, we employ an enhanced version of the GVAR methodology, initially introduced by Pesaran et al. (2004) by adding impulse saturation techniques (IIS) developed by Castle et al. (2012) and Johansen and Nielsen (2009) among other papers and applied to GVAR modelling by Ericsson and Reisman (2012) which we refer to as GVAR-IIS throughout this study. Our second objective is to show how this methodology can be useful in generating forecast fuel prices in various Brazilian regions. To achieve this, we utilize a weekly database containing data on gasoline, ethanol, and diesel prices in the Brazilian market. The study spans from May 2004 to August 2021, with the model estimation period encompassing January 2004 to December 2018 and the validation forecast period spanning from January 2019 to August 2021. Lastly, our third objective involves analyzing how the regional fuel markets in Brazil would respond to a shock in the global oil market. This investigation sheds light on the potential outcomes of such a scenario and its impact on the Brazilian fuel landscape, particularly in the context of gasoline and ethanol.

The data from fuel prices are taken from the Brazilian Ministry of Mines and Energy (Ministério de Minas e Energia, 2021) at a municipality level. We also use information collected by the Brazilian Institute of Geography and Statistics (IBGE)[2] that maps infrastructure and economic linkages across all municipalities in Brazil. This information about infrastructure is used to construct a measure of economic connections. Such a measure creates the weights used in GVAR-IIS to model interdependence across regions.

This study is divided into sections as follows: section 2 is a brief literature review on the subject; section 3 presents the methodology; section 4 presents the empirical application; section 5 reports the results of the empirical application; and section 6 contains final considerations.

[2]Initials come from "Instituto Brasileiro de Geografia e Estatística".

2 Literature Review

Countries and regions are affected by macroeconomics shocks with different intensities and at different moments in time. The 1990 Persian Gulf crisis had a global shock to the oil market that brought attention to the response of retail gasoline prices to fluctuations in world oil prices. More recently, the world was hit by the global shock brought by the COVID-Sars events. These types of shocks are international and affect economies worldwide but not necessarily at the same time. This implies that regions can react differently to them. There is also the presence of monetary policies, fiscal policies, fiscal crisis, political turmoils, and other types of region specif shocks that affect every region and affect the economy. There has, over the years, been a wealth of studies looking into a multitude of features covering the nexus of fuel prices and financial markets. This section revisits only a selection of the most important contributions offered by the previous literature.

One of the earliest studies was done by Hamilton (1983). The author analyzed how oil shocks co-moved with changes in economy-wide activity. The author described that over the period 1948-1972, the correlation between oil and United States recessions is statistically significant and nonspurious, supporting the proposition that oil shocks were a contributing factor in at least some of the United States recessions before 1972. The proposed model, though simple, illustrated that oil price movements in some manner influenced most periods of recession in the United States.

In succession, Hooker (1996) found that oil prices no longer Granger caused United States macroeconomic recessions after 1973. The author explores several potential explanations: that sample stability issues are responsible, that oil prices are endogenous, and that linear and symmetric specifications misrepresent the form of the oil price interaction. None of these hypotheses was supported by the data.

Hamilton (1996) responded to Hooker (1996) with an innovative application of econometrics that combined concepts of regime-switching with asymmetric variable decompositions. Hamilton (1996) stated that many of the quarterly oil price increases observed since 1985 were corrections to even higher oil price decreases the previous quarter. Hamilton (1996) concluded that when one looks at the net increase in oil prices over the year, the data is consistent with the historical correlation between oil shocks and recessions.

The existence of asymmetric price adjustment price was tackled by Shin (1994). The author analysed whether the perception of asymmetric price adjustment is valid. The author analyzed the issue at both the retail and the wholesale level. The empirical evidence presented in the study supports the hypothesis that the adjustment of the volume-weighted product prices to changes in crude oil prices is symmetric. In particular, the study has shown that wholesale gasoline prices adjust symmetrically to changes in crude prices. Moreover, if any asymmetry exists, it is in the opposite direction to conventional wisdom: wholesale gasoline prices fall faster than they rise. For retail gasoline prices, the findings were mixed, and the overall hypothesis that retail price adjustment is symmetrical could not be rejected by the data, particularly over longer periods.

In his turn, Borenstein et al. (1997) studied the response of gasoline prices to changes in crude oil prices. The evidence gathered by the authors supports the theory that retail gasoline prices respond more quickly to increases in crude oil prices than to decreases. An increase in oil prices might be passed along to terminal prices in the short run. Still, the shock effects would not be observable after a ten-week adjustment period. The authors show that spot prices for generic gasoline have asymmetry in responding to crude oil price changes, reflecting inventory adjustment effects. Asymmetry also appears in the response of retail prices to wholesale price changes, possibly indicating short-run market power among retailers. Borenstein et al. (1997) concluded that this result is consistent with the theoretical work of Benabou and Gertner (1993), which demonstrates that consumers may search less when the common input prices of all retailers become more variable, causing short-run decreases in the elasticity of demand that each retailer faces.

Grasso and Manera (2007) provided a detailed comparison of the three most popular models designed to describe asymmetric price behaviour. The author compared three models: the asymmetric error correction model, the autoregressive threshold error correction model and the error correction model with threshold cointegration. Each model is estimated on a common monthly data set for the gasoline markets of France, Germany, Italy, Spain and UK over the period 1985–2003. The conclusion is that all models can capture the temporal delay in the reaction of retail prices to changes in spot gasoline and crude oil prices, as well as some evidence of asymmetric behaviour.

The asymmetric response of gasoline prices to Oil price Shocks have been extensively researched. Kang et al. (2019) studied the effect of oil price shocks on the real price of gasoline with economic policy uncertainty. Kristoufek and Lunackova (2015) investigated the effects between retail gasoline and crude oil prices in a new framework of fractional integration, long-term memory and borderline (non)stationarity. Apergis and Vouzavalis (2018) looked over the asymmetric pass-through of oil prices to gasoline prices under the non-linear autoregressive distributed lags model. Sun et al. (2019) employed the threshold autoregressive interval-valued models developed recently by Sun et al. (2018) to investigate the pass-through of crude oil prices to retail gasoline prices and proposed a consistent interval-based test to detect threshold effects.

At the firm level, Broadstock et al. (2016) investigated whether firms reacts to international oil prices and if firms also react to gasoline prices, which are more directly connected to most firms costs of doing business than oil prices are. The study applied multi-factor asset pricing models to a sample of 963 Chinese firms (between 2005–2013). The author's main findings are as follows: Around 90% of Chinese firms are affected by both oil and gasoline shocks in the long run; The effects differ for price rises and price falls; The results also vary widely between industrial sectors.

Using a panel-asymmetric error correction model based on daily panel data of heterogeneous refiners, Chen and Sun (2021) explored the dynamics of China's gasoline price response to international oil market price fluctuations and the domestic price regulation. The author stated several conclusions: Firstly, China's gasoline price has an asymmetric response to international crude oil price changes. It responds on time and is closely in line with crude oil price increases but shows a lagged and long-lasting decrease response. Secondly, symmetric and asymmetric price responses to price regulation are found at the industry level as well as in refiners with different types of ownership. Specifically, state-owned refiners respond more to regulated price increases, whereas private refiners respond more to regulated price decreases. Thirdly, China's gasoline price responds symmetrically to fuel oil price changes but asymmetrically to price regulation in the context of crude oil import regulation, reflecting the distorted oil market and price response dynamics.

Regarding macroeconomic models, individual regions are interlinked through many channels when you consider a global economy. This raises several challenges in macroeconomic models that address: handling interdependence among regions, data availability, connecting the macroeconomic environment to regions, handling global shocks and regional shocks. The first is the well-known curse of dimensionality, and the latter challenge involves the detection and incorporation in the model of structural breaks.

Focusing on the curse of dimensionality, Chudik and Pesaran (2011) suggested two approaches to handle this problem: (i) shrinkage of the parameter space and (ii) shrinkage of the data. Parameter space can be shrunk by imposing a set of restrictions, which could be, for instance, obtained from a theoretical structural model, directly on the parameters, as done in the Lasso regressions proposed by Tibshirani (1996), and adaptive Lasso, where the L1 norms in the penalty are re-weighted by data-weights Zou (2006); Zou and Hastie (2005). Alternatively, one could use techniques where prior distributions are imposed on the parameters to be estimated. The second approach to mitigating the curse of dimensionality is to shrink the data along the lines of index models. Empirical evidence suggests that few dynamic factors are needed to explain the co-movement of macroeconomic variables: Stock and Watson (1999, 2002), Giannone et al. (2004) conclude that only a few, perhaps two, factors explain much of the predictable variations, while Stock and Watson (2005) estimate as much as seven factors.

The field known as Spatial Econometrics has a long tradition in Economics handling the curse of dimensionality. They developed a set of empirical that tools allow policymakers and analysts to track and anticipate the response of regions to shocks is essential. Pesaran and co-authors introduced the Global VAR model in the empirical macroeconomic literature as a model that could handle the curse of dimensionality. Pesaran et al. (2004) introduced the GVAR methodology employing a study that modelled 11 different regions using data from 1979 through 1999. This methodology was improved by Dees et al. (2007b) with the collaboration of the European Central Bank. Elhorst et al. (2018) discusses the similarities of the GVAR model and traditional spatial econometrics models. Recently the model was extended to deal with regional analysis (Pesaran et al., 2004; Chudik and Pesaran, 2011).

Forecasting regional labour markets with GVARs is undertaken in Schanne (2015) using German regional labour market data. His study had two goals: the first was to estimate the GVAR model, as proposed by Pesaran et al. (2004), and the second was to use the developed GVAR model to construct a forecast for German labour market indicators while accounting for spatial effects. The author finds that including information about labour market policies and vacancies and accounting for the lagged and contemporaneous spatial dependence can improve the forecasts relative to a simple bivariate benchmark model. Consequently, Schanne (2015) concluded that the data indicate semi-strong cross-section dependence and provide consistent results since Germany is a polycentric economy, in contrast with the United Kingdom or France (places with clearly dominant regions). The final model used in his study is a basic GVAR specification, which uses first differences without imposing cointegration relations and does not include a dominant region.

Using a global vector autoregressive (GVAR) model Konstantakis et al. (2021) studied how the Chinese economy has managed to maintain its overall economic growth, and therefore its production, throughout various crises. The GVAR model captured the complex interactions across regions and factors. The authors use the world input–output tables to serve as the tools to construct the GVAR weight matrix, as well as Node theory for selecting the dominant economies. The authors concluded that the economies of the United States and EU17 play a dominant role. In addition, the Chinese economy is unaffected, in the long run, by unanticipated shocks in the dominant economies of the United States and EU17.

The curse of dimensionality is not the only problem high dimensional models face. Another common problem is structural breaks. An unexpected change, which can happen over time or instantaneously in the population parameter, is classified as a structural break. These changes in the population parameter can affect the econometric model causing huge forecasting errors and unreliability of the model in general. There are several types od structural breaks, they can be rapid, as with crises (financial and otherwise); they can be characterized as jumps in the parameter value (as with asset-price volatility), and they can be viewed as changes in regime (as with the 2020 COVID-Sars pandemic). The structural breaks may also grow more gradually, as with some forms of innovation and globalization.

Structural instability is an important issue to be analyzed as highlighted by Castle et al. (2016). They argued that the lack of stability of coefficients frequently caused forecast failure, and therefore we must routinely test for structural stability. Thus, the detection of breaks is essential, in one hand, we have that the recent events emphasize the need for reliable "early-warning" and "early-detection" methods for assessing the state of the economy, on the other hand, the gains from modelling a structural break may be offset by imprecisely estimated break dates and post-break parameters, which, as Elliott and Müller (2006) show, is a feature of local-to-zero breaks. Therefore, whether detected or not, breaks pose great economic and statistical inference, forecasting, and policy challenges.

The first paper to enhance GVAR with impulse indicator saturation is Ericsson and Reisman (2012). They stated that automatic model selection and impulse indicator saturation contribute two valuable tools to a coherent empirical framework for detecting structural breaks such as crises, jumps, and regime changes. Ericsson and Reisman (2012) concludes that Global vector autoregressions (GVARs) have several attractive features: multiple potential channels for the international transmission of macroeconomic and financial shocks, a standardized economically appealing choice of variables for each country or region examined, systematic treatment of long-run properties through cointegration analysis, and flexible, dynamic specification through vector error correction modelling.

Our study focuses on expanding the GVAR model with impulse indicator saturation to detect breaks and shocks. We also provide an empirical analysis of the Brazilian fuel market with a forecasting exercise. Many economic shocks have hit the Brazilian economy during the last twenty years. In the eighties, Brazil was hit by the "Debt crisis" that affected many Latin American countries (Cline, 1995). In the nineties, a series of financial crisis events hit emerging markets such as Mexico, South Korea, and Russia, among others and finally, Brazil had to let their currency float facing a massive devaluation. In the first decade of the twenty-first century, oil prices rose drastically, reaching highs of \$147 in July 2008. This ended in December 2008, when the global economy entered a recession reducing the oil prices from \$147 to \$32 per barrel (World Atlas, 2021). The recession was the subprime crisis, in which Brazil was also hit with the spillovers of the situation (Kolb, 2011). At the end of the second decade of the twenty-first century, a populist policy and political turmoil took place in Brazil (Melo, 2016). The country faced one of the most profound and most prolonged recessions documented by CODACE [3]. On 8 March 2020, Saudi Arabia initiated a price war on oil with Russia, causing a 65% quarterly fall in the price of oil (Journal of Petroleum Technology, 2020). All these events turned Brazil into a unique case for understanding the effects of frequent and severe macroeconomic events on regions.

[3] CODACE is the Brazilian Business Cycle Dating Committee similar to the NBER Cycle Dating Committee. See Picchetti (2019) for a description of the methodology adopted by the Committee.

3 Methodology

This section presents and describes the Global Vector Autoregression (GVAR), which is the main methodology used in this study. It is assumed that one is familiar with the Vector Autoregression (VAR) methodology and Vector Error Correction model (VECM), as well as their terminologies. The current approach to modeling GVARs has been developed in Pesaran et al. (2004) and Dees et al. (2007b). For further research on GVARs, see Garratt et al. (2006); Pesaran and Smith (2006); Dees et al. (2007a); Vansteenkiste and Hiebert (2011); Pesaran et al. (2009); Castrén et al. (2010); Chudik and Pesaran (2011); and Smith and Galesi (2014)

3.1 Global vector autoregressive model

The GVAR methodology has several attractive features: A versatile structure for characterizing international macroeconomic and financial linkages through multiple channels, a standardized economically appealing choice of variables (both domestic and foreign) for each country or region, a systematic treatment of long-run properties through cointegration analysis, and a flexible, dynamic specification through vector error correction modelling. These features are very appealing, and they balance the roles of data and economic theory in empirical modelling.

The GVAR explicitly aims to capture international economic linkages, especially linkages between the macroeconomic and financial sides of economies. Weak exogeneity plays an essential role in allowing conditional subsystem analysis on a region-by-region basis.

The traditional GVAR model handles the curse of dimensionality by shrinking the parameter space. However, the regions/individuals of the model can be subjected to shocks, which can lead to structural breaks in the parameters of regression models, leading to substantial forecasting errors and unreliability of the model in general. A novel approach to handle unexpected shocks was proposed by Hendry (1999), which introduced the Impulse Indicator Saturation (IIS) as a test for an unknown number of breaks, occurring at unknown times, with unknown duration and magnitude. The procedure relies on adding a pulse dummy as an intervention at every observation in the sample.

The use of IIS tackles the problem of the breaks; however, it saturates the model, and the estimation is unfeasible with traditional methods. The IIS in the Autometrics routine of Doornik (2009) OxMetrics is an elegant estimation algorithm. The process utilizes many blocks, and the partitioning of the sample into blocks may vary over iterations of searches; see also Hendry and Krolzig (1999, 2001, 2005), Hoover and Perez (1999, 2005). IIS is a statistically valid procedure for integrated, cointegrated data; see Johansen and Nielsen (2009). IIS can also serve as a diagnostic statistic for many forms of misspecification.

This study proposes to extend the classical GVAR model with the inclusion of impulse-saturation dummies. The augmented GVAR model would lead to a more robust and parsimonious model, where in theory, the model will have several advantages: (i) the classical GVAR methodology eliminates the curse of dimensionality and provides a parsimonious model for each region; (ii) the dummies saturation allows the model to take into account any structural break of the population parameters or regional shock.

3.2 The Classical GVAR Model

The GVAR methodology was originally proposed by Pesaran et al. (2004) as a practical approach to construct a coherent global model. The model is usually summarized as a two-step procedure.

1. In the first step, a specific models for each region is estimated conditioned to external influences. The models are represented as VAR models, which have domestic variables and foreign variables, where the foreign variables are treated as weakly exogenous.

2. In the second stage, the VAR models of each region are grouped into one single global VAR model.

To present the classical GVAR methodology we follow the notation in Chudik and Pesaran (2016). Consider a panel of N cross-section regions (traditionally called units), each featuring k_i variables observed during the time periods $t = 1,2,\dots,T$. A set of item-specific endogenous variables are collected in a $k_i \times 1$ vector $x_{i,t}$ and let $X_t = (x'_{1,t}, x'_{2,t}, \dots, x'_{N,t})'$ denote a $k \times 1$ vector of all the variables in the panel, where $k = \sum_{i=1}^{N} k_i$. A set of foreign variables $x^*_{i,t}$ are calculated as cross-section weighted averages of foreign variables, collected in a $k^* \times 1$ vector.

$$x^*_{i,t} = \tilde{W}'_i X_t \tag{1}$$

The specific models of each region consist of a set of domestic and foreign variables, modeled as an autoregressive model with p_i lags. A set of foreign variables x^* enter the model time contemporaneously and with a number of lags up to q_i, that is:

$$x_{i,t} = a_{i,0} + a_{i,1}t + \sum_{l=1}^{p_i} \Phi_{i,l} x_{i,t-l} + \Lambda_{i,0} x^*_{i,t} + \sum_{l=1}^{q_i} \Lambda_{i,l} x^*_{i,t-l} + \epsilon_{i,t} \tag{2}$$

for $i = 1, 2, \dots, N$, where $\Phi_{i,l}$ is an array of lag coefficients for lag l associated with domestic variables and $\Lambda_{i,l}$ is an array of lag coefficients for lag l associated with foreign variables; and $\epsilon_{i,t}$ is as an error vector.

Let $Z_{i,t} = (x'_{i,t}, x^{*\prime}_{i,t})'$ be $k_i + k^*$ dimensional vector of domestic and region-specific foreign variables included in the submodel of region i and rewrite (2) as

$$\mathbf{A}_{i,0} Z_{i,t} = a_{i,0} + a_{i,1} t + \sum_{\ell=1}^{p} \mathbf{H}_{i,\ell} Z_{i,t-\ell} + \epsilon_{it} \tag{3}$$

where $\mathbf{A}_{i,0} = (\mathbf{I}_{k_i}, -\Lambda_{i,0})\,, \mathbf{H}_{i\ell} = (\Phi_{i,\ell}, \Lambda_{i,\ell})$ for $\ell = 1, 2, \ldots, p$. Also, $p = \max_i(p_i, q_i)$, and define $\Phi_{i,\ell} = 0$ for $\ell > p_i$, and similarly $\Lambda_{i,\ell} = 0$ for $\ell > q_i$. Individual region-models in (3) can be equivalently written in the form of error-correction representation.

$$\mathbf{A}_{i,0} \Delta Z_{i,t} = \sum_{\ell=1}^{p} \mathbf{A}_{i,\ell} \Delta Z_{i,t-\ell} - \mathbf{\Pi}_i Z_{i,t-1} + \epsilon_{i,t} \tag{4}$$

where $\Delta = 1 - L$ is the usual first difference operator, and $\mathbf{\Pi}_i = \mathbf{A}_{i,0} - \sum_{\ell=1}^{p} \mathbf{H}_{i,\ell}$, and $\mathbf{A}_{i,\ell} = -(\mathbf{H}_{i,\ell+1} + \mathbf{H}_{i,\ell+2} + \cdots + \mathbf{H}_{i,\ell+p})$. Also, we omit the constant term in the model for readability.

The data shrinkage comes from the use of cross-section weighted averages of foreign variables, given by (1), which solves the dimensionality problem. The cross-section averages $x^*_{i,t}$ is treated as weakly exogenous and therefore (2) can be estimated consistently (notice that the weak exogeneity assumption is testable). The estimation of regional models in (2), which allows for cointegration within and across regions (via the star variables), is the first step of the GVAR approach. The second step of the GVAR approach consists of stacking estimated regional models to form one large global VAR model.

3.3 Impulse indicator saturation

Until equation (4), all the model definition follows the classical GVAR model. The impulse indicator saturation (IIS) uses zero–one impulse indicator dummies to analyze properties of a model. The model in equation (4) is then augmented by a set of IIS such that there are T such dummies, one for each observation in the sample. The model is also augmented by a set of centered seasonal dummies. Therefore we rewrite equation (4) as follows:

$$\mathbf{A}_{i,0} \Delta Z_{i,t} = \sum_{\ell=1}^{p} \mathbf{A}_{i,\ell} \Delta Z_{i,t-\ell} - \mathbf{\Pi}_i Z_{i,t-1} + \mathbf{U}_i \zeta + \mathbf{IIS}_i \Psi + \epsilon_{i,t} \tag{5}$$

were ζ is a matrix of centered seasonal dummies and $\mathbf{U}_i$ are the set of coefficients associated to the centered seasonal dummies; Ψ is a $(k_i + k^*) \times T$ matrix that represents the Impulse Indicator dummies and $\mathbf{IIS}_i$ is the set of coefficients associated to them. Using the $(k_i + k^*) \times k$ dimensional 'link' matrices $W_i = (E'_i, \tilde{W}'_i)'$, where E_i is $k \times k_i$ selection matrix that select $x_{i,t}$, namely $x_{i,t} = E'_i x_t$, and $\tilde{W}_i$ is the weight matrix introduced in (1) to define region-specific foreign star variables. We have

$$Z_{i,t} = \begin{bmatrix} x'_{i,t} & x^{*}_{i,t}{}' \end{bmatrix}' = W_i X_t \tag{6}$$

Notice that region-specific models allow for cointegration both among domestic variables as well as between domestic and foreign (star) variables. In particular, assuming $Z_{i,t}$ is integrated of order 1, the rank of Π_i specifies the number of cointegrating relationships that exist among the domestic and region-specific foreign variables in $Z_{i,t}$; and Π_i can be decomposed as $\Pi_i = \alpha_i \beta'_i$. Therefore, using (6) in (5), and decomposing Π_i yields

$$\mathbf{A}_{i,0} W_i \Delta X_t = \sum_{\ell=1}^{p} \mathbf{A}_{i,\ell} W_i \Delta X_{t-\ell} - \alpha_i \beta'_i Z_{i,t-1} + \mathbf{U}_i \zeta + \mathbf{IIS}_i \Psi + \epsilon_{i,t} \tag{7}$$

At the core of the GVAR approach are small-scale region-specific conditional models that can be estimated separately. These individual regions models explain the domestic variables of a given economy, $x_{i,t}$, conditional on region-specific cross-section averages of foreign variables. However, the inclusion of all T dummies makes the model model estimation unfeasible by conventional methods.

3.4 Global model with IIS

The second step of the GVAR approach consists of stacking estimated region models to form one large global VAR model. Therefore, stacking the equations (7) for each region and defining the following set of matrices:

$$\mathbf{G}_0 = \begin{bmatrix} \mathbf{A}_{1,0} W_1 \\ \vdots \\ \mathbf{A}_{n,0} W_n \end{bmatrix} \qquad \mathbf{G}_l = \begin{bmatrix} \mathbf{A}_{1,l} W_1 \\ \vdots \\ \mathbf{A}_{n,l} W_n \end{bmatrix} \qquad \mathbf{G}_\alpha = \begin{bmatrix} -\alpha_1 \beta'_1 W_1 \\ \vdots \\ -\alpha_n \beta'_n W_n \end{bmatrix} \qquad \mathbf{U}_{stk} = \begin{bmatrix} \mathbf{U}_i \\ \vdots \\ \mathbf{U}_n \end{bmatrix}$$

$$\mathbf{IIS}_{stk} = \begin{bmatrix} \mathbf{IIS}_1 \\ \vdots \\ \mathbf{IIS}_n \end{bmatrix} \qquad u_t = \begin{bmatrix} \epsilon_{1,t} \\ \vdots \\ \epsilon_{n,t} \end{bmatrix}$$

We then have the following equation:

$$\mathbf{G}_0 \Delta X_t = \sum_{\ell=1}^{p} \mathbf{G}_\ell \Delta X_{t-\ell} + \mathbf{G}_\alpha X_{t-1} + \mathbf{U}_{stk} \zeta + \mathbf{IIS}_{stk} \Psi + u_t \tag{8}$$

If matrix $\mathbf{G}_0$ is invertible, then by multiplying (8) by $\mathbf{G}_0^{-1}$ from the left we obtain the solution to the GVAR model in the error-correction representation:

$$\Delta X_t = \sum_{\ell=1}^{p} \mathbf{G}_0^{-1}\mathbf{G}_\ell \Delta X_{t-\ell} + \mathbf{G}_0^{-1}\mathbf{G}_\alpha X_{t-1} + \mathbf{G}_0^{-1}\mathbf{U}_{stk}\zeta + \mathbf{G}_0^{-1}\mathbf{IIS}_{stk}\Psi + \mathbf{G}_0^{-1}u_t \tag{9}$$

If matrix G_0 is not invertible, then the system (8) is undetermined, and to additional equations are required for X_t to be uniquely determined. Pesaran and Chudik (2011) show that these additional equations can be specified in the form of VAR models in cross-sectional averages of $x_{i,t}$

3.4.1 Common Variables

As in the case of the classical GVAR model, common variables can be introduced in the region models, either as observed common factors or in the form of dominant variables as defined in Chudik and Pesaran (2013). Let ω_t be vector of $m_w \times 1$ observed common factors variables. In this case, the conditional region models are included in the model as ω_t and its lagged values; therefore, after piling the equation (8) is rewritten as:

$$\mathbf{G}_0\Delta X_t = \sum_{\ell=1}^{p} \mathbf{G}_\ell \Delta X_{t-\ell} + \mathbf{G}_\alpha X_{t-1} + \mathbf{U}_{stk}\zeta + \mathbf{IIS}_{stk}\Psi + \mathbf{D}_{stk,0}\Delta w_t + \sum_{\ell=1}^{s_i} \mathbf{D}_{stk,\ell}\Delta w_{t-\ell} + u_t \tag{10}$$

for $i = 1, 2, \ldots, N$. where $\mathbf{D}_{stk,\ell}$ are the coefficients associated with the common factors $\omega_{t-\ell}$ for $\ell = 1, \ldots, s_i$. Both types of variables (common variables ω and cross-section averages $x^*_{i,t}$) can be treated as weakly exogenous for the purpose of estimation. As noted, the weak exogeneity assumption is testable. The marginal model for the dominant variables can be estimated with or without the feedback effects from x_t. In the latter case, we have the following marginal model:

$$\Delta w_t = -\alpha_w \beta'_w w_{t-1} + \sum_{l=1}^{p_w} H_{w,l}\Delta w_{t-\ell} + \mathbf{U}_w\zeta + \mathbf{IIS}_w\Psi + \eta_t \tag{11}$$

Feedback effects from the variables in the GVAR model back to the dominant variables via cross-section averages, therefore (11) can be augmented by lags of $X^*_{w,t} = \bar{W}_w X_t$, where $\bar{W}$ dimensional weight matrix defining global cross-section averages. Assuming there is no cointegration among the common variables, ω_t, and the cross-section averages, $X^*_{w,t-\ell}$ (notice that this condition is also testable), them the marginal model for the dominant variables equation with the feedback becomes:

$$\Delta w_t = -\alpha_w \beta_w' w_{t-1} + \sum_{\ell=1}^{p_w} H_{w,\ell} \Delta w_{t-\ell} + \sum_{\ell=1}^{q_w} B_{w,\ell} \bar{W}_w \Delta X_{t-\ell} + \mathbf{U}_w \zeta + \mathbf{IIS}_w \Psi + \eta_t \tag{12}$$

different lag orders for the dominant variables (p_w) and cross-section averages (q_w) could be considered. Note that contemporaneous values of star variables do not feature in (12). Conditional models (10) and the marginal model (12) can be combined and solved as a complete global VAR model in the usual way.

$$\begin{aligned} \underbrace{\begin{bmatrix} \mathbf{I}_w & \mathbf{0} \\ -\mathbf{D}_{stk,0} & \mathbf{G}_0 \end{bmatrix}}_{\mathbf{G}_{y,0}} \begin{bmatrix} \Delta w_t \\ \Delta X_t \end{bmatrix} = \sum_{\ell=1}^{\tilde{p}} \underbrace{\begin{bmatrix} H_{w,\ell} & B_{w,\ell} \bar{W}_w \\ \mathbf{D}_{stk,\ell} & \mathbf{G}_\ell \end{bmatrix}}_{\mathbf{G}_{y,\ell}} \begin{bmatrix} \Delta w_{t-\ell} \\ \Delta X_{t-\ell} \end{bmatrix} + \underbrace{\begin{bmatrix} \mathbf{U}_w & \mathbf{IIS}_w \\ \mathbf{U}_{stk} & \mathbf{IIS}_{stk} \end{bmatrix}}_{\mathbf{C}} \begin{bmatrix} \zeta \\ \Psi \end{bmatrix} + \\ + \underbrace{\begin{bmatrix} -\alpha_w \beta_w' & 0 \\ 0 & \mathbf{G}_\alpha \end{bmatrix}}_{\mathbf{L}} \begin{bmatrix} w_{t-1} \\ X_{t-1} \end{bmatrix} + \begin{bmatrix} \eta_t \\ u_t \end{bmatrix} \end{aligned} \tag{13}$$

where $\mathbf{I}_w$ is an identity matrix of dimension m_w, $\mathbf{0}$ is a zero matrix of dimensions $m_w \times k$, $\tilde{p} = \max_i(p, p_w, q_w, s_i)$, and define $H_{w,\ell} = 0$ for $\ell > p_w$, $B_{w,\ell} = 0$ for $\ell > q_w$, $\mathbf{D}_{stk,\ell} = 0$ for $\ell > s_i$, $\mathbf{G}_\ell = 0$ for $\ell > p$.

Let $y_t' = \begin{bmatrix} w_t & X_t \end{bmatrix}$, $\Omega' = \begin{bmatrix} \zeta & \Psi \end{bmatrix}$ and $\varepsilon_t' = \begin{bmatrix} \eta_t & u_t \end{bmatrix}$, then equation (13) can be rewritten as

$$\mathbf{G}_{y,0} \Delta y_t = \sum_{\ell=1}^{\tilde{p}} \mathbf{G}_{y,\ell} \Delta y_{t-\ell} + \mathbf{C}\Omega + \mathbf{L} y_{t-1} + \varepsilon_t \tag{14}$$

if $\mathbf{G}_{y,0}$ is invertable, then

$$\Delta y_t = \sum_{\ell=1}^{\tilde{p}} \mathbf{G}_{y,0}^{-1} \mathbf{G}_{y,\ell} \Delta y_{t-\ell} + \mathbf{G}_{y,0}^{-1} \mathbf{C}\Omega + \mathbf{G}_{y,0}^{-1} \mathbf{L} y_{t-1} + \mathbf{G}_{y,0}^{-1} \varepsilon_t \tag{15}$$

3.5 GVAR and Weak exogeneity

As noted earlier the GVAR approach builds on separate estimation of region-specific VAR models assuming that the foreign variables can be treated as weakly exogenous. The concept of weak exogeneity in a system of an integrated of order one variables is also closely related to the notions of "long-run causality" and "long-run forcing" discussed by Granger and Lin (1995) and Pesaran et al. (2000). The GVAR-IIS model is build on the same principals of the classical GVAR and must follow the same considerations for the estimation of region-specific VAR models, and therefore the assumption that the foreign variables can be treated as weakly exogenous is essential.

The assumption of weak exogeneity can be easily tested as outlined in section 7.1 of Pesaran et al. (2004), and typically is not rejected, when the economy under consideration is small relative to the rest of the world and the weights used in the construction of the star variables are granular. As stated in Pesaran et al. (2004), we can check the weak exogeneity by testing the joint significance of the estimated error-correction terms.

4 Empirical Application

In this section, we illustrate our methodology by modeling the Brazilian fuel market's dynamics from May 2004 to August 2021. Brazil, the largest country in South America and Latin America, encompasses roughly 8.5 million square kilometers and has a population exceeding 210 million (as of 2019). With a robust Gross Domestic Product (GDP) of 1.84 trillion USD, the Brazilian fuel market stands out for its diversity and distinctive attributes.

Brazil has been at the forefront of innovation in the transportation sector, particularly with the development and widespread adoption of flex-fuel vehicles. These adaptable vehicles can seamlessly operate on various fuel blends, including gasoline, ethanol, or any combination thereof. Remarkably, in 2020, 83% of the total market comprised flex-fuel vehicles (ANFAVEA, 2021).

Moreover, Brazil has implemented regulatory mandates that require the incorporation of biodiesel into diesel fuel, aligning with the country's commitment to cleaner energy sources. Given these factors, the pricing dynamics in the Brazilian fuel market can be susceptible to fluctuations influenced by various elements, including global oil prices, currency exchange rates, and government policies related to fuel subsidies and taxation.

4.1 Model

This study will use the GVAR methodology with the IIS augmentation (GVAR-IIS), as described in the previous section, to evaluate the Brazilian fuel market model.

Let i be the aggregation units of several municipalities[4]. The Brazilian fuel market has several products: hydrous ethanol, regular gasoline, additive gasoline, liquefied petroleum gas, compressed natural gas, diesel oil, and diesel oil s10. However, the three major fuels used in the market are diesel oil, regular gasoline, and hydrous ethanol (Centro Brasileiro de Infraestrutura, 2019). Therefore, the specific model for each region will consist of the prices for diesel oil, regular gasoline, and hydrous ethanol. Consequently, in period t for each region i a vector $x_{i,t}$ can be defined as

[4]A full description of each municipality and its aggregation can be found in the appendix

$$x_{i,t} = \begin{bmatrix} Ethanol_{i,t} \\ Gasoline_{i,t} \\ Diesel_{i,t} \end{bmatrix}, \tag{16}$$

where $Ethanol_{i,t}$ is the natural logarithm of the price of hydrous ethanol for region i at time t, $Gasoline_{i,t}$ is the natural logarithm of the price of regular gasoline for region i at time t, and $Diesel_{i,t}$ is the natural logarithm of the price of regular diesel oil for region i at time t.

Macroeconomic variables are treated as common variables and modelled in a dominant unit. Therefore, the dominant unit of the GVAR-IIS model will contain the log level of the Brent oil price and the log level of the nominal exchange rate between Brazilian Real and the United States Dollar (BRL/USD). Let ω_t be a vector of macroeconomic variables, then the dominant unit will be a model as defined in (12), and ω_t expressed as

$$\omega_t = \begin{bmatrix} Brent_t \\ ExchangeRate_t \end{bmatrix} \tag{17}$$

4.1.1 Weight matrix

The first step in the GVAR modeling exercise is to construct the foreign region-specific ("starred") variables from the domestic variables using (1). In Pesaran et al. (2004) the weight matrix is defined by country-specific trade weights based on the United Nations Direction of Trade Statistics.

According to di Mauro and Pesaran (2013) the weight matrix W_i is meant to capture the importance of all other regions for the region i. In other words, the matrix W_i plays a key role in linking up the models of the different regions together and shows the degree to which one region depends on the remaining regions. Therefore, to determine the weight matrix, it is necessary to decide on the importance of each region for every other region.

Since we will apply the model to the fuel market, to determine the importance of a region to another is to determine the consumers prefer to acquire convenience goods, in this case, fuel, given a set of constraints which in the case of spatial competition consist in a connection between the regions (road, railroad, etc.). We will assume that the products are homogeneous among regions. The connections between regions in the modern world can be any means of connection that allow the consumer to reach (by car, by train, by plane, etc.) the seller in another region. Regions with many connections between them are more important than regions with fewer connections. Consequently, the importance of region j to region i is modelled as a function of the connections between them.

Therefore, for the weights, we use a matrix that is based on will be based on the number of connections between the regions. Chudik and Pesaran (2016) argues that the construction of cross-sectional averages only need to satisfy some granularity conditions, and for large N asymptotics one might as well use equal weights, namely, replace all cross-sectional averages by simple averages. The author also argues that when the number of regions is moderate and spillover effects could also be significant, it is advisable to use weights that also capture political and cultural linkages across regions.

Therefore, the weight matrix is a function of the number of connections between two regions. Let I be the set of all the regions in our model, and $\mathcal{I}_{i,j}$ be an indicator function where it assumes the value one if the region i have any connection with the region j; otherwise, it will assume the value zero. We define the weight vector as $\tilde{W}_i = (a_{i,1}, a_{i,2}, \ldots, a_{i,N})$, then each element $a_{i,j}$ of the weight matrix vector is defined as:

$$a_{i,j} = \frac{\mathcal{I}_{i,j} w_j}{\sum_{j \in I} \mathcal{I}_{i,j} w_j} \tag{18}$$

were $\mathcal{I}_{i,i} = 0$ and w_j is a weighting factor associated with region j that measures its importance. This weighting factor is determined to take into account the economic influence and market size. In our analysis, the weighting factor w_j is chosen to be the gross domestic product per capita of region j.

4.2 Database Description

The following section provides a brief description of the data used in this study as well as their sources and frequency. Table 1 contains a detailed description of the data sources used in the empirical analysis.

Tabela 1: Data sources and characteristics

Data	Level of data	Frequency	Period
Brazilian fuel prices[1]	municipal	Weekly	2004-05-09 to 2021-08-29
Brazilian Exchange Rate[2]	federal	Daily	2004-05-09 to 2022-01-20
Brent Crude Oil prices[3]	global	Daily	2004-03-01 to 2021-11-15
Geographic regions[4]	municipal	N/A	as reported in 2015
Connection between regions[4]	municipal	N/A	as of 2007

Source: elaborated by the author

[1] Acquired from National Agency of Petroleum, Natural Gas and Biofuels

[2] Acquired from Brazilian Central Bank

[3] Acquired from ICE Intercontinental Exchange - Provided by Bloomberg terminal

[4] Acquired from Brazilian Institute of Geography and Statistics

The data for the Brazilian fuel prices were obtained from the National Agency of Petroleum, Natural Gas and Biofuels (ANP from the Portuguese *Agência Nacional do Petróleo, Gás Natural e Biocombustíveis*). The database is available by municipality region and was grouped into mesoregions[5] taking the average price for each region and

[5]Mesoregions are geographical subdivisions that fall between the broader regional level and the more localized microregion or district level.

each fuel. A detailed description of each region group is in the appendix. Also, missing values were estimated using a local level space state model.

The data for the Brazilian Exchange Rate was obtained from the Brazilian Central Bank (Brazilian Real to United States Dollar). The series is reported at a daily frequency and were transformed to a weekly frequency using the average value of the daily data. In its turn, Brent Crude Oil prices are obtained from the ICE Intercontinental Exchange. This series also has a daily frequency and were also transformed to a weekly frequency using the average value of the daily data.

Finally, the geographic regions and municipalities information as well as the the number of connections between regions were obtained from IBGE (2008). This work aims at determining the influence of the cities in Brazil. The study not only analyzed regular transport links, such as those that go to the urban centers, but also the main destinations of the residents to obtain products and services, such as general goods, education, airport travel, and health services. The acquisition of agricultural inputs and the destination of the agricultural products were also analyzed in the study. The main result of the paper is a "Brazilian urban network" with a mapping of the connections for all 5,564 Brazilian municipalities. This "Brazilian urban network" contains the hierarchy of the urban network and the regions of influence of each urban center. Information on the total population, their designation codes, were also obtained from IBGE[6] as reported in 2015.

4.3 Descriptive statistics

Table 2 shows descriptive statistics for the Brent Crude Oil, nominal exchange rate, as well as the fuel prices (Hydrous Ethanol, Diesel Oil, Regular Gasoline) for the regions that contain the cities of São Paulo, Rio de Janeiro, Belo Horizonte, Brasilia, and Salvador - five of the most important cities in Brazil. The descriptive statistics for all the regions can be found in the appendices.

[6]IBGE from the Portuguese *Instituto Brasileiro de Geografia e Estatística*

Tabela 2: Descriptive statistics

Region	Series	Mean	Median	Max	Min	Std_Dev	Skewness	Kurtosis
Sao Paulo	Hydrous Ethanol [1]	1.99	1.89	4.44	0.74	0.70	0.78	3.46
Sao Paulo	Diesel Oil [1]	2.48	2.12	4.59	1.39	0.70	0.82	2.86
Sao Paulo	Regular Gasoline [1]	3.05	2.70	5.71	1.84	0.81	1.03	3.35
Rio de Janeiro	Hydrous Ethanol	2.54	2.27	5.44	1.04	0.91	0.84	3.12
Rio de Janeiro	Diesel Oil	2.52	2.12	4.57	1.36	0.75	0.69	2.39
Rio de Janeiro	Regular Gasoline	3.36	2.92	6.49	1.98	1.00	0.99	3.10
Belo Horizonte	Hydrous Ethanol	2.29	2.19	4.69	1.15	0.65	0.95	3.88
Belo Horizonte	Diesel Oil	2.52	2.19	4.65	1.34	0.77	0.70	2.45
Belo Horizonte	Regular Gasoline	3.22	2.87	6.19	1.89	0.95	0.97	3.02
Dist. Federal	Hydrous Ethanol	2.49	2.27	5.41	1.22	0.76	0.92	3.75
Dist. Federal	Diesel Oil	2.58	2.13	4.73	1.44	0.78	0.68	2.32
Dist. Federal	Regular Gasoline	3.24	2.88	6.48	1.94	0.84	1.14	4.08
Salvador	Hydrous Ethanol	2.42	2.16	4.97	1.27	0.76	0.99	3.62
Salvador	Diesel Oil	2.47	2.10	4.64	1.31	0.72	0.77	2.74
Salvador	Regular Gasoline	3.24	2.81	5.96	1.92	0.85	1.05	3.42
	Brent Crude Oil [USD][2]	73.61	67.75	143.05	21.61	25.23	0.49	2.19
	Exchange rate [BRL/USD][3]	2.78	2.34	5.85	1.55	1.06	1.04	3.28

Source: elaborated by the author

Weekly data from May 9th, 2004 to September 4th, 2021

[1] Value in Brazilian Reais per liter of the fuel.

[2] Value for the Brent Crude Oil prices in United States Dollar.

[3] Value for the nominal exchange rate (Brazilian Real per United States Dollar)

4.4 Model configuration

The GVAR-IIS model allows for several configuration schemes. One objective of this study is to evaluate the performance of the IIS procedure in the case of model misspecification. To assess the issue of misspecification several models are analyzed: (i) the model is evaluated with or without the ISS procedure, which means the GVAR-IIS or the classical GVAR; (ii) the model can be evaluated with the long-run equilibrium estimate or with an overestimation of the long-run relationships, in our models the long run estimation is determined by the Johansen method and a higher rank of an extra cointegration relationship is also utilized; (iii) the model can be evaluated with seasonal dummies or not; (iv) lastly the IIS procedure can be allowed to remove the seasonal dummies or not. The table (3) has a description of all the models.

Model	**IIS**	**Cointegration**	**Dummies**	**Rem. Dummy**	**Exogenous**
M1	ON	Higher rank	ON	No	Brent
M2	OFF	Normal rank	OFF	No	Brent
M3	OFF	Higher rank	OFF	No	Brent
M4	OFF	Normal rank	ON	No	Brent
M5	OFF	Higher rank	ON	No	Brent
M6	ON	Normal rank	OFF	No	Brent
M7	ON	Normal rank	ON	No	Brent
M8	ON	Normal rank	ON	Yes	Brent
M9	ON	Higher rank	ON	No	Brent and Exc.Rate
M10	OFF	Normal rank	OFF	No	Brent and Exc.Rate
M11	OFF	Higher rank	OFF	No	Brent and Exc.Rate
M12	OFF	Normal rank	ON	No	Brent and Exc.Rate
M13	OFF	Higher rank	ON	No	Brent and Exc.Rate
M14	ON	Normal rank	OFF	No	Brent and Exc.Rate
M15	ON	Normal rank	ON	No	Brent and Exc.Rate
M16	ON	Normal rank	ON	Yes	Brent and Exc.Rate

Tabela 3: Forecasting Models

All models are estimated with 8 lags and the use of Autometrics in the GVAR-IIS depends on a p-value setting for the variable selection, this value is chosen a *priori*. All models that have the IIS procedure turned on have the p-value set at 10^{-5}.

A total of 110 regional models were estimated, 84 models specified with no cointegration relationship, 24 models specified with one cointegration relationship, and 2 models specified with two cointegration relationships. The procedure developed by Johansen (1992) was used to validate the weak exogeneity of the foreign variables. In a total of 108 models, the weakly exogenous can not be rejected, and in 2 models the foreign variables could not be accepted as weakly exogenous.

Taking into account structural breaks in the estimation process is one of the significant advantages of the GVAR-IIS approach, as well as taking into account misspecification models. A total of 18 models using different configurations were estimated: 2 VECM models taken to be the benchmark models, 8 models with the IIS procedure, and 8 models without the IIS procedure (classical GVAR setting). All models with the impulse saturation procedure were able to produce more accurate forecasts when compared to the benchmarks and the classical GVAR models. Even in the case of misspecification, the GVAR-IIS model can make a more reliable prediction. Using the RMSE metric, the best predictive model (M8) is the GVAR-IIS model with Brent Crude Oil as an exogenous variable.

As a benchmark comparison we consider a Vector of Error Correction Model (VECM), as it is one of the most popular models (Grasso and Manera, 2007). The VECM model is also estimated with an 8 lag, and the cointegration vector is estimated using the Johansen procedure. Two error correction models are estimated, the first considers the Brent Crude Oil as the only exogenous variable, the second model considers the Brent Crude Oil as well as the nominal exchange rate as exogenous variables.

4.4.1 Forecasting evaluation

When it comes to forecasting evaluation, a standard methodology is to reserve part of the sample data for use as a comparison sample in which the actual values and forecasted values are compared. The estimation window uses data from May 2004 to December 2018. The data from January 1st 2019, to August 29th, 2021 is set aside for forecast evaluation. The usual metric used to measure the differences between actual and predicted values by a model or an estimator is the root-mean-square error (RMSE). The metric can evaluate a specific series (diesel oil, regular gasoline, and hydrous ethanol) prediction or can evaluate all predictions, independent of the series. Let i be a specific series then the RMSE for a specific series can be defined as:

$$\text{RMSE}_i = \sqrt{\frac{\sum_{t=1}^{T} \left(\hat{y}_{i,t} - y_{i,t}\right)^2}{T}} \tag{19}$$

where $\hat{y}_{i,t}$ is the forecast value of series i at time t and $y_{i,t}$ is the actual value of series i at time t. The total RMSE can be defined as

$$\text{RMSE}_{Total} = \sqrt{\frac{\sum_{i=1}^{N} \sum_{t=1}^{T} \left(\hat{y}_{i,t} - y_{i,t}\right)^2}{NT}} \tag{20}$$

5 Results[7]

Modelling a system such as the Brazilian fuel market at the regional level is not an easy task, and it is undoubtedly subject to considerable uncertainties. There are many choices

[7]All models were coded by the author and estimated using OxMetrics version 8.00 64-bit edition, and result analysis was done in R version 4.1.0.

to consider for each region model. For instance, one might consider the variables to be included in each regional model, the lag order of the domestic variables, and the lag order of the foreign variables; one also might consider the number of cointegrating relations and whether to impose long-run or short-run restrictions on the parameters.

As explained earlier, the GVAR and GVAR-IIS models comprise several specific models solved in a two-step process. Each regional model uses the data for diesel oil, regular gasoline, and hydrous ethanol and suffers the influence of the foreign variables using the weight matrix defined previously. Also, all regions are affected by the Brent Crude Oil and nominal exchange rate.

At the time of this analysis, Brazil had 5612 municipalities that are grouped in 110 mesoregions. As stated in (18), each region has a weighting vector that is used to determine the influences of the foreign variables.

5.1 Cointegration

The GVAR-IIS and the classical GVAR allow for cointegration among regional and foreign variables[8]. All individual regional models had the deterministic components treated as an unrestricted intercept and no trend in the error correction models.

Out of the 110 estimated models, 84 were specified with no cointegration relationship, 24 were specified with one cointegration relationship, and 2 were set with two cointegration relationships.

5.2 Weak exogeneity

The estimation of region-specific VAR models assumes that the foreign variables can be treated as weakly exogenous. This is a common requirement between the classical GVAR and the GVAR-IIS. Using the procedure developed by Johansen (1992) the weak exogeneity is tested for the 110 models. In a total of 108 models, the test of weakly exogenous can not be rejected, and in 2 models the foreign variables could not be accepted as weakly exogenous. We accepted the weakly exogenous conditioning since these regions are within the 5% expected rejection threshold. It is worth noting that, according to Pesaran et al. (2004), even if the weak exogeneity assumption is rejected, one could still obtain consistent estimates of the parameters of the GVAR model in two steps under certain conditions.

[8]The cointegration was investigated using the Cointegration Analysis of Time Series by Jurgen A. Doornik and Katarina Juselius, available in the Oxmetrics version 8.0

5.3 Forecast analysis

The ability to consider structural breaks in the estimation process is one of the major advantages of the GVAR-IIS approach. The account of structural breaks should provide robust forecasts. In this section, we investigate how the forecasting performance of GVAR-IIS models compares with other forecasting models, mainly the classical GVAR and the VECM models. Table 4 show the results of the forecast exercise using the RMSE metric.

Tabela 4: Root mean squared errors

Model	**RMSE Total**	**MSE Ethanol**	**MSE Diesel**	**MSE Gasoline**
$M8_{IIS}$	1.64×10^{-2}	1.97×10^{-2}	1.53×10^{-2}	1.36×10^{-2}
$M6_{IIS}$	1.64×10^{-2}	1.96×10^{-2}	1.55×10^{-2}	1.37×10^{-2}
$M16_{IIS}$	1.65×10^{-2}	1.97×10^{-2}	1.56×10^{-2}	1.38×10^{-2}
$M14_{IIS}$	1.66×10^{-2}	1.97×10^{-2}	1.56×10^{-2}	1.38×10^{-2}
$M7_{IIS}$	1.67×10^{-2}	1.98×10^{-2}	1.58×10^{-2}	1.38×10^{-2}
$M15_{IIS}$	1.67×10^{-2}	1.99×10^{-2}	1.58×10^{-2}	1.38×10^{-2}
$M1_{IIS}$	1.70×10^{-2}	2.05×10^{-2}	1.60×10^{-2}	1.40×10^{-2}
$M9_{IIS}$	1.71×10^{-2}	2.07×10^{-2}	1.60×10^{-2}	1.40×10^{-2}
VECM2[1]	1.74×10^{-2}	2.06×10^{-2}	1.65×10^{-2}	1.44×10^{-2}
VECM	1.78×10^{-2}	2.10×10^{-2}	1.70×10^{-2}	1.48×10^{-2}
M12	2.01×10^{-2}	2.26×10^{-2}	2.08×10^{-2}	1.66×10^{-2}
M4	2.02×10^{-2}	2.25×10^{-2}	2.09×10^{-2}	1.67×10^{-2}
M10	2.04×10^{-2}	2.22×10^{-2}	2.18×10^{-2}	1.69×10^{-2}
M2	2.05×10^{-2}	2.22×10^{-2}	2.19×10^{-2}	1.70×10^{-2}
M13	2.11×10^{-2}	2.49×10^{-2}	2.05×10^{-2}	1.71×10^{-2}
M5	2.12×10^{-2}	2.49×10^{-2}	2.07×10^{-2}	1.72×10^{-2}
M11	2.17×10^{-2}	2.52×10^{-2}	2.14×10^{-2}	1.77×10^{-2}
M3	2.17×10^{-2}	2.52×10^{-2}	2.15×10^{-2}	1.78×10^{-2}

Source: elaborated by the author

[1] VECM2 refers to the model with two exogenous variables (Brent Crude Oil and nominal exchange rate)

[IIS] Model with IIS procedure

From the results presented at table 4 it is possible to verify that the forecasts produced by model M8 have the lowest root mean squared error in the overall metric. Recall that model M8 has the IIS procedure, seasonal dummies and only the Brent Crude Oil as an exogenous variable. Model M8 also produces the best forecast values for diesel and regular gasoline series. As for the Hydrous Ethanol series model, M6 has the lowest root mean squared error. Overall the models with the IIS procedure can produce better forecasts than the classical GVAR model and better results than the benchmarks (VECM and VECM2).

Likewise, the models can be evaluated considering each region. In particular, for the region that contains the municipality of São Paulo, one of the most important in the country, the best models for the series of diesel oil, hydrous ethanol, and regular gasoline

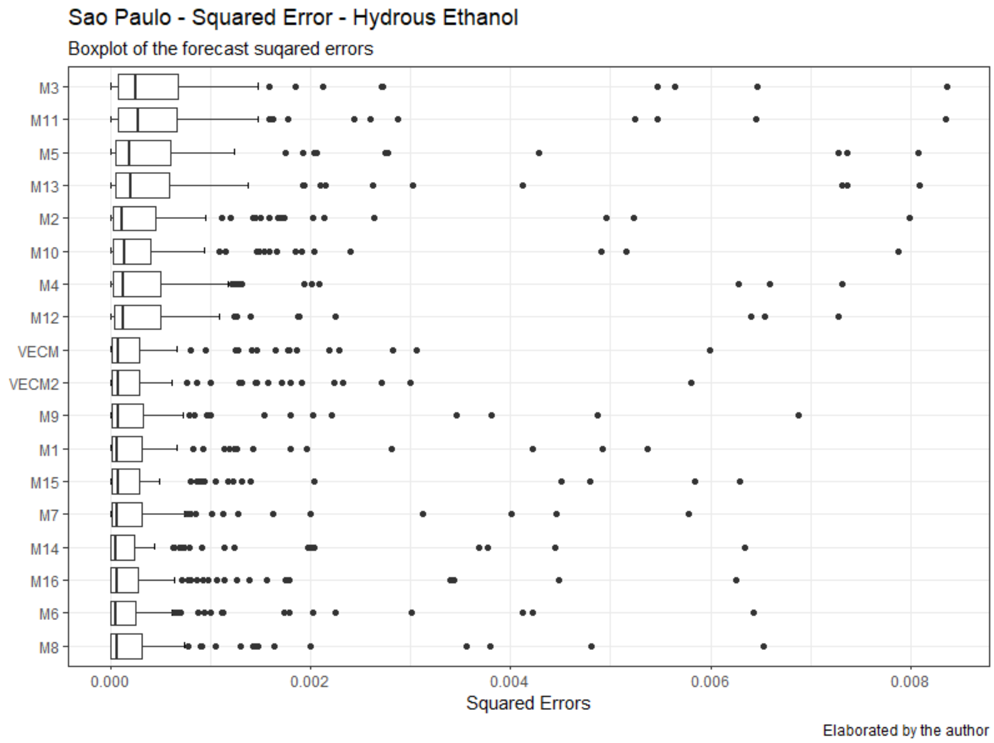


Figura 1: Boxplot of the squared errors (region: São Paulo; fuel: Hydrous Ethanol)

are the models M8, M7 and M15 (all models were the IIS procedure is turned on). Figure (1), Figure (2), and Figure (3) show a boxplot for the squared errors of the models. We can notice that, in general, models with the forecasts of the GVAR-IIS models have a lower interquartile range, indicating that these models have lower variance in the estimates.

5.4 Model Comparison

Although the root mean squared errors are a useful metric and extensively used in the literature, the metric does not consider if two competing models can be regarded as equals in predictive power. In other words, if the difference in two models is statistically significant. In order to determine whether forecasts are significantly different, the Diebold-Mariano test (Diebold and Mariano, 1995) is applied. The test originally proposed by Diebold and Mariano (1995) considers a sample path of loss differentials. Under the assumption that the loss differential is a covariance stationary series, the sample average converges asymptotically to a normal distribution.

Table (5) present the p-values of the Diebold-Mariano test for predictive accuracy. Only the models labelled: M8, M6, M16, M14, M7, M15, M1, M9, VECM2, and VECM are presented. A table with a complete comparison between models can be found in the appendix. According to the Diebold-Mariano test, model M8 is considered the best model for predictive accuracy. Also, all GVAR-IIS models have better predictive accuracy than the classical GVAR and the VECM benchmark. It is worth mentioning that both VECM

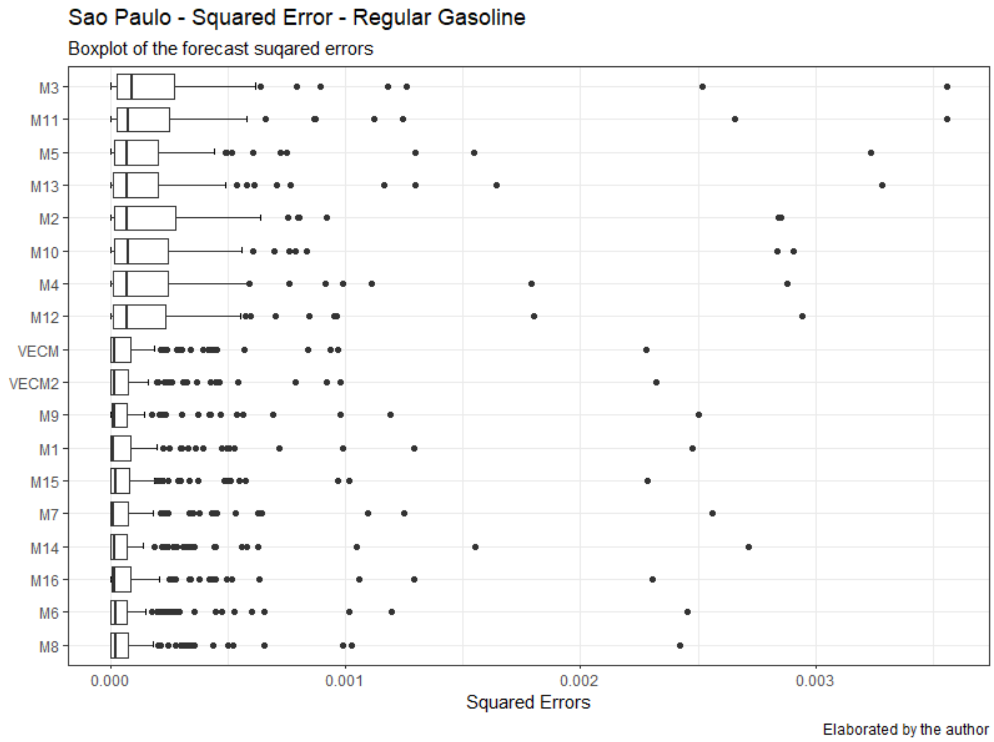


Figura 2: Boxplot of the squared errors (region: São Paulo; fuel: Regular Gasoline)

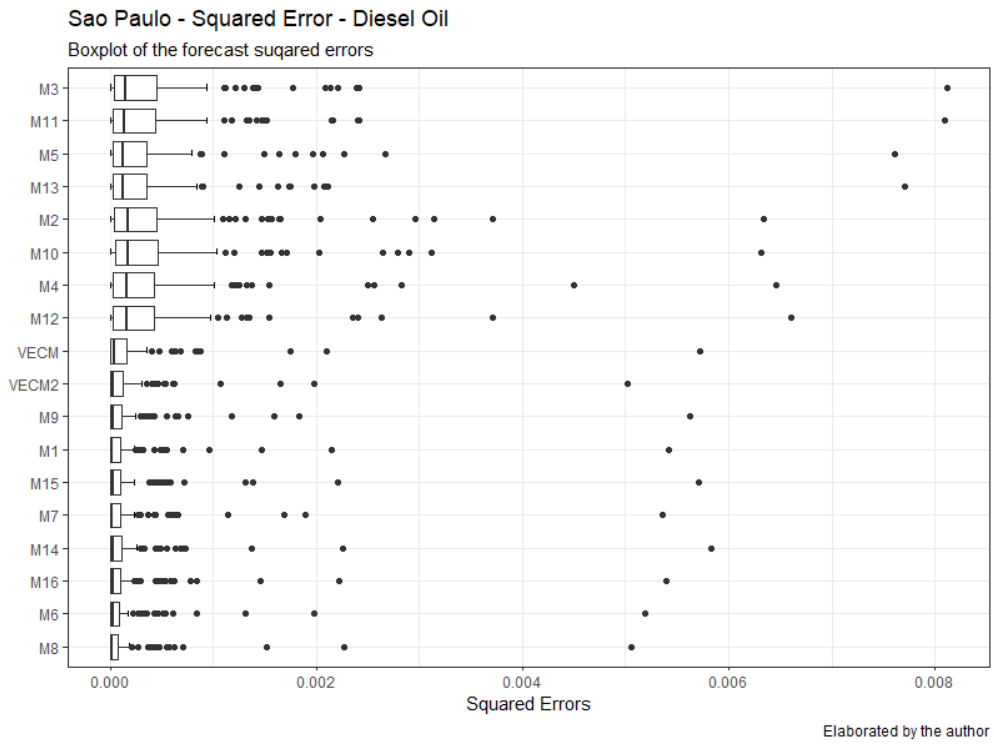


Figura 3: Boxplot of the squared errors (region: São Paulo; fuel: Diesel Oil)

models, considered to be the benchmark, has better predictive accuracy when compared to the classical GVAR.

Tabela 5: P-values for the Diebold-Mariano test for predictive accuracy

	M8	**M6**	**M16**	**M14**	**M7**	**M15**	**M1**	**M9**	**VECM2**	**VECM**
M8		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
M6	0.000		1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000
M16	0.000	0.000		0.000	1.000	1.000	1.000	1.000	1.000	1.000
M14	0.000	0.006	1.000		1.000	1.000	1.000	1.000	1.000	1.000
M7	0.000	0.000	0.000	0.000		0.994	1.000	1.000	1.000	1.000
M15	0.000	0.000	0.000	0.000	0.006		1.000	1.000	1.000	1.000
M1	0.000	0.000	0.000	0.000	0.000	0.000		1.000	1.000	1.000
M9	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.999	1.000
VECM2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001		1.000
VECM	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	

Diebold-Mariano test with one step ahead forecast
Loss function is the squared-error loss
Alternative hypothesis: row model error is greater than column model error

5.5 Impulse Response Function

So far, all the GVAR-IIS models have similar results; however, the results for the GVAR-IIS model M16, which is the model with the lowest RMSE with both the Brent Crude Oil and nominal exchange rate as exogenous variables, will be present in this section.

Using the GVAR methodology, it is possible to estimate the dynamic response of the fuel market for each region, considering not only the interdependence of each region but also the effects of the exogenous variables. With this in mind, we investigate the fuel market's response to a shock in the Brent Crude Oil market.

The impulse response describes the system's reaction as a function of time for each fuel: diesel oil, regular gasoline, and hydrous ethanol. More generally, an impulse response is the system's response to some external change. Figure (4) presents the impulse response function for the regions containing the municipalities of São Paulo, Rio de Janeiro, and Belo Horizonte. Notice that, for these regions, the regular gasoline peaks at approximately five weeks, showing no significant changes after eight weeks. As for hydrous ethanol and diesel oil, they have similar behaviours for the selected regions. It is worth mentioning that the diesel oil presents a more prominent correction in the eighth period reaching stability twelve weeks after the initial shock.

Using the GVAR model, it is possible to estimate the dynamic response of a single region as well as the response of the system as a whole; figure (5) presents the system response for the whole market. The shock reveals a significant positive effect on the hydrous ethanol in the first week, presenting a positive response to the shock. At the same time, gasoline and Diesel do not offer significant changes. This fast response can be attributed to the freedom of hydrous ethanol prices. The regular gasoline, in its turn, reveals a positive effect predominantly in the central and northern regions of the country. After five weeks, all fuels contained the effects of the shock showing a substantial improvement in prices, reaching the peak in price changes. After eight weeks, the shock effects are dissipated for regular gasoline and hydrous ethanol. However, diesel oil has a correction effect in most of the regions in the country. After ten weeks, the shock effect is dissipated in the diesel oil series.

5.6 Individual model forecasts

On Monday 20th April, West Texas Intermediate (WTI) Crude Oil, the benchmark for US oil, fell as low as minus 37.63 USD a barrel. On that same day, the Brent Crude Oil, the standard measure used by Europe and the rest of the world, was trading based on June contracts and reached a value of 35.75 USD. For the first time in history, the future price of oil was negative. Although the economic community has provided numerous explanations for this event, we can evaluate how the GVAR-IIS model handle the forecast in the presence of such a shock. The GVAR-IIS model M16 is the model with the best predictive

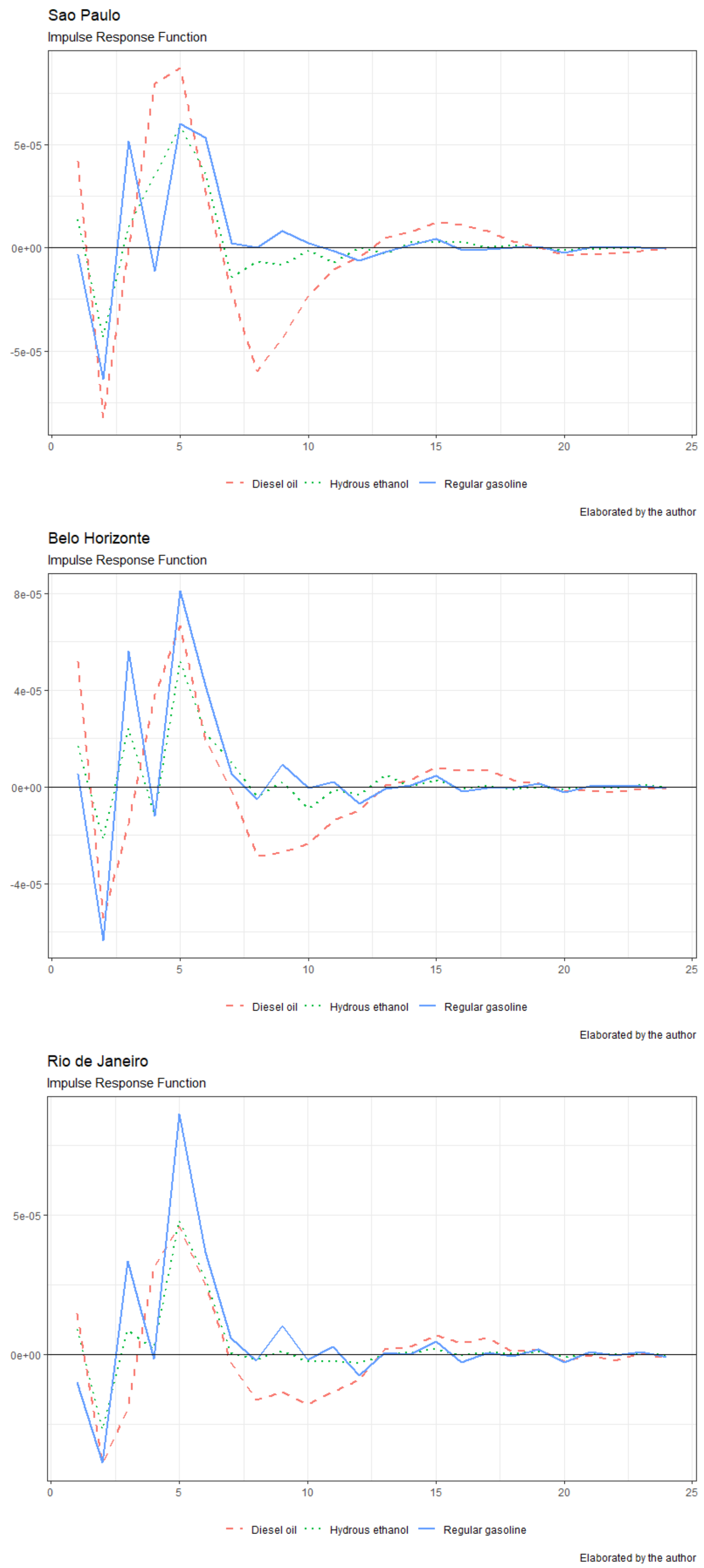
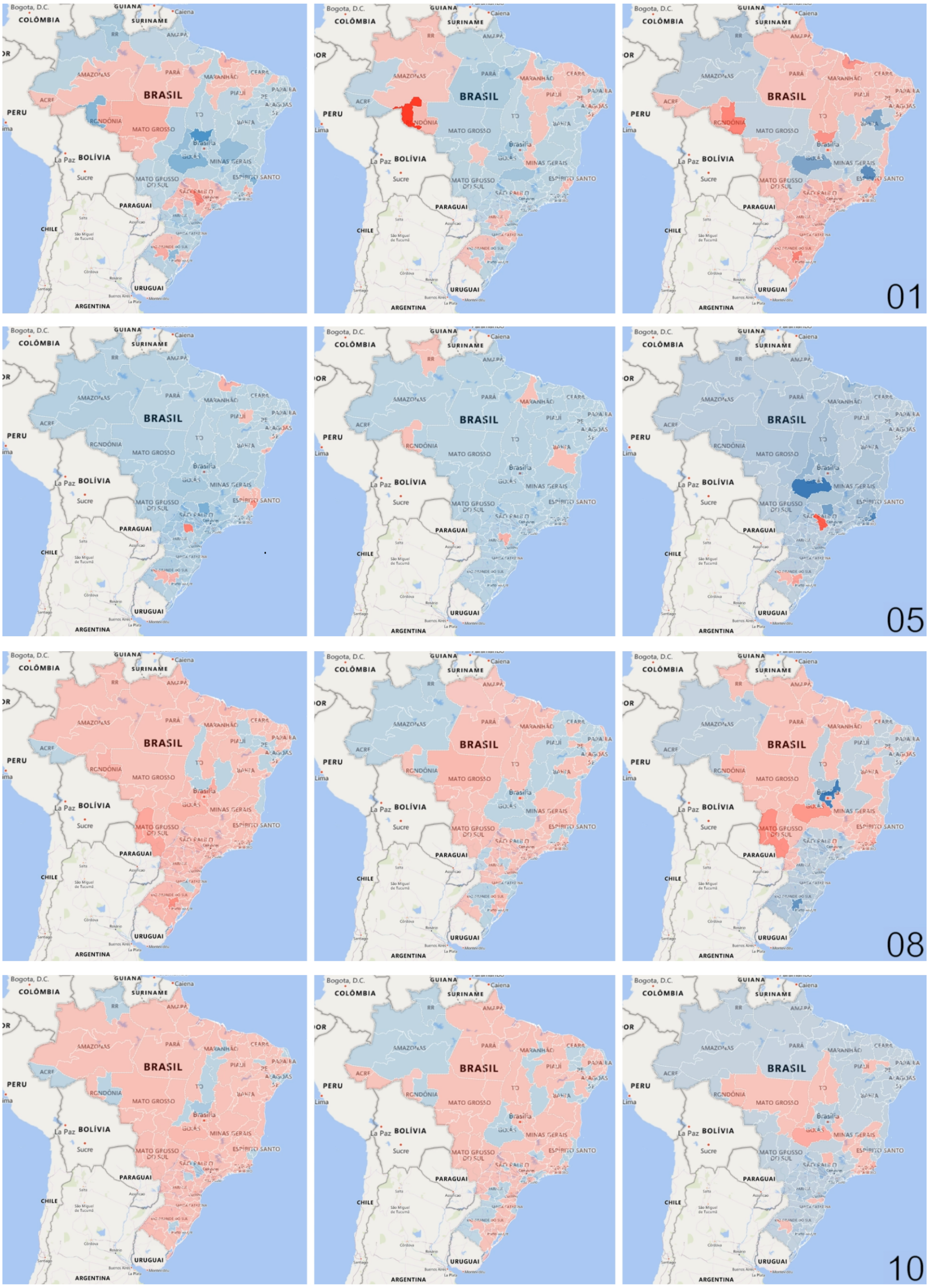


Figura 4: Impulse response function for the region of São Paulo, Belo Horizonte and Rio de Janeiro



Columns (left to right): Diesel oil, hydrous ethanol, regular gasoline.
Rows (top to bottom): 1st period, 5th period, 8th period and 10th period
(blue) increase in prices; (red) decrease in prices.

Figura 5: Impulse response function for all regions (periods 1, 5, 8 and 10)

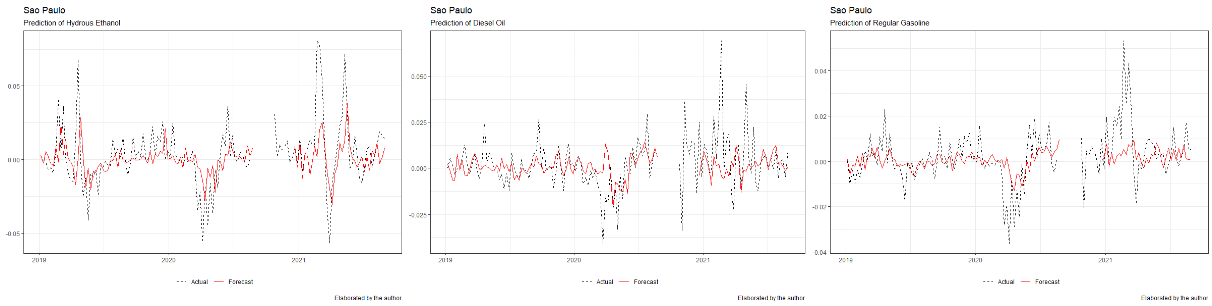


Figura 6: GVAR-IIS M16 forecasts of hydrous ethanol, diesel oil, and regular gasoline for the region of São Paulo

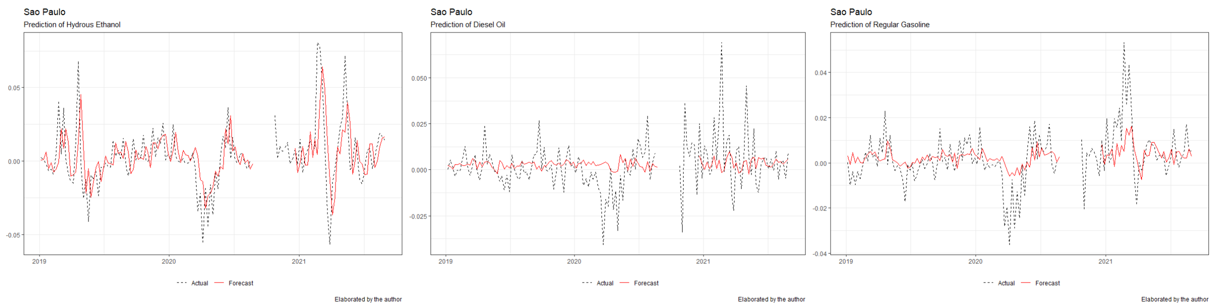


Figura 7: VECM forecasts of hydrous ethanol, diesel oil, and regular gasoline for the region of São Paulo

accuracy with Brent Crude Oil and nominal exchange rate as exogenous variables and therefore will is the model chosen to be compared against the benchmark.

The GVAR-IIS model can produce solid forecasts of the diesel oil, hydrous ethanol, and regular gasoline in each region, considering not only the effects of the macro variables such as nominal exchange rate and the Brent Crude Oil prices but also the interdependence of each region. We will focus on the region containing the São Paulo municipality since it is the most populated in the country.

Figure (6) shows the forecast predictions for the GVAR-IIS model M16, and figure (7) shows the forecast predictions for the VECM model benchmark (with Brent Crude Oil and nominal exchange rate as exogenous variables). The benchmark model does not capture the effects of the oil shock in the diesel oil and has a low impact on regular gasoline. In comparison, the GVAR-IIS model M16 can capture some of the effects of the shock over the diesel oil and has more suitable forecasts for the regular gasoline. Therefore the GVAR-IIS forecast has a good performance in the presence of an unexpected shock, adapting more quickly to the shock. It is worth mentioning that the COVID restrictions caused a nine-week period of missing data for all regions, which caused a loss of predictions in the sample.

6 Conclusion

One of the goals of them paper was to present an extended version of classical Global Vector Autoregression proposed by Pesaran et al. (2004). The classical model was augmented with Impulse–Indicator Saturation introduced by Hendry (1999). The second goal was to show how the methodology performed in an empirical exercise.

With this in mind, the study discusses an empirical application focusing on forecasting the Brazilian fuel market. The classical GVAR methodology, developed by Pesaran et al. (2004), plays a leading role in the modelling of the fuel market since it accounts for the interdependence and connections between economies and regions, and the IIS procedure (Hendry, 1999) can account for structural breaks in the series and therefore produce more reliable forecasts. As in the classical GVAR model, the assumptions of weak exogeneity are still required for the GVAR-IIS model.

A classical Diebold-Mariano test is performed to evaluate the model's predictive accuracy. The best model is confirmed as the GVAR-IIS model with Brent Crude Oil as an exogenous variable. It is worth mentioning that the classical GVAR forecasts perform worse than the benchmark models.

The response of the system to a global oil price shock is examined using impulse response functions. The study delves into the behaviour of three key fuel series: diesel oil, regular gasoline, and hydrous ethanol.

In the initial week, diesel oil exhibited positive responses in a significant portion of the country. By the fifth week, all three series reached their peak, demonstrating positive reactions to the oil price shock. However, after eight weeks, the effects of the shock start to dissipate for regular gasoline and hydrous ethanol.

Notably, diesel oil sustains its response to the shock for an extended period, with the duration ranging from 10 to 12 weeks depending on the region. This implies a more protracted impact of the shock on the diesel market compared to regular gasoline and hydrous ethanol.

Finally, the model is evaluated during the oil shock of 2020, where oil prices suffered a significant negative shock. During the period of the shock, the benchmark model is unable to capture the effects of the oil shock in the diesel oil series and shows a low impact on the regular gasoline series. In turn, the GVAR-IIS model M16 can capture some of the effects of the shock over diesel oil and possess more suitable predictions for regular gasoline. Therefore, the GVAR-IIS model is more robust in the presence of an unexpected shock, adapting more quickly.

In conclusion, our analysis indicates that the GVAR-IIS methodology can deal with the curse of dimensionality, misspecification and structural breaks. The GVAR-IIS can produce reliable forecasts in the case of model misspecification and has a better performance when compared to classical GVAR or VECM models.

Referências Bibliográficas

ANFAVEA, A. N. d. F. d. V. A. (2010). Anúario da industria automobilistica brasileira 2010. *Brazilian Automotive Industry Yearbook*.

ANFAVEA, A. N. d. F. d. V. A. (2021). Anúario da industria automobilistica brasileira 2021. *Brazilian Automotive Industry Yearbook*.

Apergis, N. and Vouzavalis, G. (2018). Asymmetric pass through of oil prices to gasoline prices: Evidence from a new country sample. *Energy Policy*, 114(September 2017):519–528.

Balke, N. S., Brown, S. P. a., and Yucel, M. K. (1998). Crude oil and gasoline prices: An asymmetric relationship? *Economic Review (07321414)*, page 2.

Benabou, R. and Gertner, R. (1993). Search with learning from prices does increased inflationary uncertainty lead to higher markups. *Review of Economic Studies*, 60(1):69–94.

Bennett, J. K., Owyang, M. T., and Vermann, E. K. (2021). Regional gasoline price dynamics. *Federal Reserve Bank of St. Louis Review*, 103(3):289–314.

Borenstein, S., Colin Cameron, A., and Gilbert, R. (1997). Do gasoline prices respond asymmetrically to crude oil price changes? *Quarterly Journal of Economics*, 112(1):304–339.

Broadstock, D. C., Fan, Y., Ji, Q., and Zhang, D. (2016). Shocks and stocks: A bottom-up assessment of the relationship between oil prices, gasoline prices and the returns of Chinese firms. *Energy Journal*, 37(15):55–86.

Castle, J. L., Clements, M. P., and Hendry, D. F. (2016). An Overview of Forecasting Facing Breaks. *Journal of Business Cycle Research*, 12(1):3–23.

Castle, J. L., Doornik, J. A., and Hendry, D. F. (2012). Model selection when there are multiple breaks. *Journal of Econometrics*, 169(2):239–246.

Castrén, O., Dées, S., and Zaher, F. (2010). Stress-testing euro area corporate default probabilities using a global macroeconomic model. *Journal of Financial Stability*, 6(2):64–78.

Centro Brasileiro de Infraestrutura (2019). Como é feita a distribuição de combustíveis?

Chen, H. and Sun, Z. (2021). International crude oil price, regulation and asymmetric response of China's gasoline price. *Energy Economics*, 94:105049.

Chudik, A. and Pesaran, M. H. (2011). Infinite-dimensional VARs and factor models. *Journal of Econometrics*, 163(1):4–22.

Chudik, A. and Pesaran, M. H. (2013). Econometric Analysis of High Dimensional VARs Featuring a Dominant Unit. *Econometric Reviews*, 32(5-6):592–649.

Chudik, A. and Pesaran, M. H. (2016). Theory and practice of gvar modelling. *Journal of Economic Surveys*, 30(1):165–197.

Cline, W. R. (1995). International debt reexamined. *Peterson Institute Press: All Books*.

Dees, S., Holly, S., Pesaran, M. H., and Smith, L. V. (2007a). Long Run Macroeconomic Relations in the Global Economy. *Economics: The Open-Access, Open-Assessment E-Journal*, 1(2007-3):1.

Dees, S., Mauro, F. d., Pesaran, M. H., and Smith, L. V. (2007b). Exploring the international linkages of the euro area: a global VAR analysis. *Journal of Applied Econometrics*, 22(1):1–38.

di Mauro, F. and Pesaran, M. H. (2013). *The GVAR Handbook: Structure and Applications of a Macro Model of the Global Economy for Policy Analysis*. Oxford University Press.

Diebold, F. X. and Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, 13(3):253–263.

Doan, T., Litterman, R., and Sims, C. (1984). Forecasting And Conditional Projection Using Realistic Prior Distributions. *Econometric Reviews*, 3(1):1–100.

Doornik, J. A. (2009). Autometrics. In Jeniffer L. Castle and Neil Shephard, editor, *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*, chapter 4, pages 88–121. Department of Economics, University of Oxford.

Elhorst, J. P., Gross, M., and Tereanu, E. (2018). Spillovers in space and time: where spatial econometrics and Global VAR models meet. *ECB Working Paper*, 0(2134).

Elliott, G. and Müller, U. K. (2006). Efficient tests for general persistent time variation in regression coefficients. *Review of Economic Studies*, 73(4):907–940.

Ericsson, N. R. and Reisman, E. L. (2012). Evaluating a Global Vector Autoregression for Forecasting. *International Advances in Economic Research*, 18(3):247–258.

Garratt, A., Lee, K., Pesaran, M. H., and Shin, Y. (2006). *Global and National Macroeconometric Modelling: A Long-Run Structural Approach*. Oxford University Press.

Giannone, D., Reichlin, L., and Sala, L. (2004). Monetary policy in real time. *NBER Macroeconomics Annual*, 19:161–200.

Granger, C. W. and Lin, J. L. (1995). Causality in the long run. *Econometric Theory*, 11(3):530–536.

Grasso, M. and Manera, M. (2007). Asymmetric error correction models for the oil-gasoline price relationship. *Energy Policy*, 35(1):156–177.

Hamilton, J. D. (1983). Oil and the Macroeconomy since World War II. *Journal of Political Economy*, 91(2):228–248.

Hamilton, J. D. (1996). This is what happened to the oil price - Macroeconomy relationship. *Journal of Monetary Economics*, 38(2):215–220.

Hendry, D. F. (1999). An Econometric Analysis of US Food Expenditure. *Methodology and Tacit Knowledge: Two Experiments in Econometrics*, pages 341–361.

Hendry, D. F. and Krolzig, H. (1999). Improving on 'Data mining reconsidered' by K.D. Hoover and S.J. Perez. *The Econometrics Journal*, 2(2):202–219.

Hendry, D. F. and Krolzig, H.-M. (2001). Computer Automation of General-to-Specific Model Selection Procedures. *Journal of Economic Dynamics and Control*, 25:831–866.

Hendry, D. F. and Krolzig, H. M. (2005). The properties of automatic gets modelling. *Economic Journal*, 115(502).

Hooker, M. A. (1996). What happened to the oil price-macroeconomy relationship? *Journal of Monetary Economics*, 38:195–213.

Hoover, K. D. and Perez, S. J. (1999). Data mining reconsidered: encompassing and the general-to-specific approach to specification search. *The Econometrics Journal*, 2(2):167–191.

Hoover, K. D. and Perez, S. J. (2005). Truth and Robustness in Cross-country Growth Regressions. *SSRN Electronic Journal*.

IBGE (2008). *Regiões de influência das cidades 2007*, volume 1. Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro.

Johansen, S. (1992). Cointegration in partial systems and the efficiency of single-equation analysis. *Journal of Econometrics*, 52(3):389–402.

Johansen, S. and Nielsen, B. (2009). An Analysis of the Indicator Saturation Estimator as a Robust Regression Estimator. *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*.

Journal of Petroleum Technology (2020). OPEC+ Moves To End Price War With 9.7

Million B/D Cut.

Kang, W., de Gracia, F. P., and Ratti, R. A. (2019). The asymmetric response of gasoline prices to oil price shocks and policy uncertainty. *Energy Economics*, 77:66–79.

Kasibhatta, S. (2019). Alcohol fuels as an alternative fuels-bringing new heights in sustainability. *Alcohol Fuels-Current Technologies and Future Prospect.*

Kolb, R. W. (2011). *Financial contagion: the viral threat to the wealth of nations*, volume 604. John Wiley & Sons.

Konstantakis, K. N., Michaelides, P. G., Chatzieleftheriou, L., and Prelorentzos, A. N. (2021). Crisis and the Chinese miracle : A network—GVAR model. *Bulletin of Economic Research*, 1(November):1–22.

Kristoufek, L. and Lunackova, P. (2015). Rockets and feathers meet Joseph: Reinvestigating the oil-gasoline asymmetry on the international markets. *Energy Economics*, 49:1–8.

Melo, M. A. (2016). Crisis and Integrity in Brazil. *Journal of Democracy*, 27(2):50–65.

Ministério de Minas e Energia (2021). Série histórica do Levantamento de Preços. https://www.gov.br/anp/pt-br/assuntos/precos-e-defesa-da-concorrencia/precos/precos-revenda-e-de-distribuicao-combustiveis/serie-historica-do-levantamento-de-precos.

Pesaran, M., Shin, Y., and Smith, R. J. (2000). Structural analysis of vector error correction models with exogenous I(1) variables. *Journal of Econometrics*, 97(2):293–343.

Pesaran, M. H. and Chudik, A. (2011). Aggregation in Large Dynamic Panels. *Federal Reserve Bank of Dallas, Globalization and Monetary Policy Institute Working Papers*, 2011(101).

Pesaran, M. H., Schuermann, T., and Smith, L. V. (2009). Forecasting economic and financial variables with global VARs. *International Journal of Forecasting.*

Pesaran, M. H., Schuermann, T., and Weiner, S. M. (2004). Modeling Regional Interdependencies Using a Global Error-Correcting Macroeconometric Model. *Journal of Business & Economic Statistics*, 22(2):129–162.

Pesaran, M. H. and Smith, R. (2006). Macroeconometric modelling with a global perspective. *Manchester School*, 74(SUPPL. 1):24–49.

Picchetti, P. (2019). Brazilian Business Cycles as Characterized by CODACE. In *Business Cycles in BRICS*, pages 331–335. Springer.

Schanne, N. (2015). A Global Vector Autoregression (GVAR) model for regional labour markets and its forecasting performance with leading indicators in Germany. *IAB-Discussion Paper*, 1(13).

Shin, D. (1994). Do product prices respond symmetrically to changes in crude prices? *OPEC Review*, 18(2):137–157.

Smith, L. and Galesi, A. (2014). GVAR Toolbox 2.0. *University of Cambridge: Judge Business School.*

Stock, J. H. and Watson, M. W. (1999). Forecasting inflation. *Journal of Monetary Economics*, 44(2):293–335.

Stock, J. H. and Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics*, 20(2):147.

Stock, J. H. and Watson, M. W. (2005). Implications of Dynamic Factor Models for VAR Analysis. *NBER Working Paper Series*, 11467:1–67.

Sun, Y., Han, A., Hong, Y., and Wang, S. (2018). Threshold autoregressive models for interval-valued time series data. *Journal of Econometrics*, 206(2):414–446.

Sun, Y., Zhang, X., Hong, Y., and Wang, S. (2019). Asymmetric pass-through of oil prices to gasoline prices with interval time series modelling. *Energy Economics*, 78:165–173.

Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288.

Vansteenkiste, I. and Hiebert, P. (2011). Do house price developments spillover across euro area countries? Evidence from a global VAR. *Journal of Housing Economics*, 20(4):299–314.

World Atlas (2021). 5 Worst Energy Crises Of All Time.

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American Statistical Association*, 101(476):1418–1429.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 67(2):301–320.

Tabela 6: Descriptive statistics

Region	Serie	Mean	Median	Max	Min	Std_Dev	Skewness	Kurtosis
1	Hydrous Ethanol	2.66	2.44	5.22	1.40	0.80	0.67	2.73
1	Diesel Oil	2.73	2.32	4.86	1.51	0.74	0.66	2.46
1	Regular Gasoline	3.34	2.99	6.08	2.20	0.80	1.02	3.43
2	Hydrous Ethanol	2.80	2.46	5.59	1.57	0.90	0.80	2.68
2	Diesel Oil	2.76	2.37	5.02	1.56	0.76	0.77	2.77
2	Regular Gasoline	3.49	3.02	6.29	2.36	0.85	1.14	3.61
3	Hydrous Ethanol	3.02	2.72	5.96	1.80	0.85	1.05	3.89
3	Diesel Oil	3.15	2.69	5.85	1.71	0.94	0.82	2.76
3	Regular Gasoline	3.74	3.24	6.75	2.39	0.93	1.09	3.51
4	Hydrous Ethanol	2.72	2.57	4.95	1.54	0.75	0.52	2.37
4	Diesel Oil	2.75	2.37	5.00	1.52	0.77	0.63	2.42
4	Regular Gasoline	3.41	3.02	6.17	2.04	0.86	0.78	2.67
5	Hydrous Ethanol	2.84	2.60	5.10	1.30	0.76	0.56	2.35
5	Diesel Oil	2.81	2.49	4.85	1.58	0.66	0.81	3.35
5	Regular Gasoline	3.30	2.96	5.70	2.00	0.69	0.96	3.78
6	Hydrous Ethanol	2.95	2.60	5.64	1.74	0.78	0.88	2.94
6	Diesel Oil	2.72	2.30	5.05	1.49	0.81	0.70	2.45
6	Regular Gasoline	3.58	3.10	6.44	2.39	0.87	0.90	2.85
7	Hydrous Ethanol	2.80	2.44	5.48	1.69	0.81	0.89	3.01
7	Diesel Oil	2.64	2.22	4.73	1.42	0.78	0.63	2.18
7	Regular Gasoline	3.33	2.88	6.02	2.22	0.80	1.03	3.23
8	Hydrous Ethanol	2.53	2.24	5.25	1.22	0.87	0.87	2.93
8	Diesel Oil	2.54	2.16	4.70	1.42	0.75	0.79	2.62
8	Regular Gasoline	3.37	2.95	6.21	2.08	0.88	1.09	3.39
9	Hydrous Ethanol	2.53	2.27	5.27	1.24	0.87	0.82	2.95
9	Diesel Oil	2.51	2.11	4.59	1.48	0.69	0.89	2.92
9	Regular Gasoline	3.33	2.92	6.22	2.10	0.86	1.15	3.58
110	Hydrous Ethanol	2.90	2.49	6.35	1.71	0.88	0.86	3.01
110	Diesel Oil	2.80	2.36	5.03	1.52	0.82	0.73	2.54
110	Regular Gasoline	3.36	2.99	5.94	2.25	0.73	1.10	3.87
10	Hydrous Ethanol	2.56	2.27	5.18	1.28	0.82	0.78	2.73
10	Diesel Oil	2.51	2.15	4.66	1.36	0.73	0.75	2.66
10	Regular Gasoline	3.15	2.79	5.90	1.96	0.77	1.12	3.78
11	Hydrous Ethanol	2.60	2.24	5.19	1.57	0.80	0.79	2.73
11	Diesel Oil	2.52	2.11	4.69	1.35	0.76	0.78	2.55
11	Regular Gasoline	3.25	2.79	6.06	2.01	0.82	1.03	3.38

Source: elaborated by the author
Weekly data from May 9th, 2004 to September 4th, 2021
All values are in Brazilian Reais per liter of fuel. Except for the “Brent crude oil” and “Exchange rate”
[1] Value for the Brent crude oil prices in United States Dollar.
[2] Value for the nominal exchange rate (Brazilian Real per United States Dollar)

Tabela 7: Descriptive statistics

Region	Serie	Mean	Median	Max	Min	Std_Dev	Skewness	Kurtosis
12	Hydrous Ethanol	2.65	2.40	3.93	1.46	0.68	0.53	1.78
12	Diesel Oil	2.51	2.13	3.98	1.35	0.69	0.52	1.87
12	Regular Gasoline	3.25	2.85	4.84	2.04	0.74	0.72	2.15
13	Hydrous Ethanol	2.70	2.41	5.08	1.59	0.78	0.85	2.74
13	Diesel Oil	2.52	2.14	4.64	1.36	0.76	0.85	2.84
13	Regular Gasoline	3.35	2.93	5.91	2.16	0.80	1.18	3.85
14	Hydrous Ethanol	2.69	2.43	5.41	1.57	0.74	1.09	4.13
14	Diesel Oil	2.60	2.16	5.08	1.41	0.79	0.88	2.93
14	Regular Gasoline	3.29	2.78	6.50	2.10	0.92	1.31	3.99
15	Hydrous Ethanol	2.57	2.33	5.46	1.46	0.71	1.13	4.48
15	Diesel Oil	2.55	2.13	4.88	1.33	0.78	0.81	2.80
15	Regular Gasoline	3.17	2.71	6.50	2.04	0.89	1.31	4.12
16	Hydrous Ethanol	2.67	2.39	3.90	1.66	0.62	0.48	1.80
16	Diesel Oil	2.51	2.14	3.90	1.38	0.68	0.52	1.91
16	Regular Gasoline	3.25	2.76	5.02	2.16	0.80	0.80	2.15
17	Hydrous Ethanol	2.72	2.38	5.59	1.66	0.83	1.05	3.48
17	Diesel Oil	2.60	2.19	4.91	1.51	0.78	0.82	2.71
17	Regular Gasoline	3.42	2.95	6.21	2.22	0.84	1.10	3.46
18	Hydrous Ethanol	2.68	2.34	5.55	1.52	0.82	0.94	3.11
18	Diesel Oil	2.61	2.19	4.90	1.59	0.77	0.76	2.39
18	Regular Gasoline	3.41	2.90	6.21	2.19	0.86	1.01	2.97
19	Hydrous Ethanol	2.60	2.33	5.52	1.41	0.83	0.94	3.34
19	Diesel Oil	2.63	2.20	4.82	1.52	0.78	0.76	2.44
19	Regular Gasoline	3.35	2.86	6.14	2.24	0.86	1.08	3.27
20	Hydrous Ethanol	2.57	2.27	5.50	1.27	0.85	0.86	3.12
20	Diesel Oil	2.58	2.14	4.91	1.48	0.81	0.71	2.36
20	Regular Gasoline	3.29	2.82	6.09	2.10	0.88	0.97	2.96
21	Hydrous Ethanol	2.67	2.34	5.45	1.55	0.82	1.03	3.61
21	Diesel Oil	2.62	2.19	4.94	1.57	0.78	0.84	2.67
21	Regular Gasoline	3.38	2.86	6.39	2.09	0.89	1.03	3.21
22	Hydrous Ethanol	2.53	2.29	5.64	1.18	0.85	1.00	3.80
22	Diesel Oil	2.51	2.10	4.64	1.30	0.76	0.72	2.49
22	Regular Gasoline	3.29	2.87	6.47	1.91	0.89	1.08	3.57
23	Hydrous Ethanol	2.58	2.35	4.87	1.42	0.79	0.89	3.23
23	Diesel Oil	2.53	2.11	4.64	1.36	0.79	0.83	2.72
23	Regular Gasoline	3.25	2.82	5.81	1.98	0.87	1.08	3.52

Source: elaborated by the author
Weekly data from May 9th, 2004 to September 4th, 2021
All values are in Brazilian Reais per liter of fuel. Except for the “Brent crude oil” and “Exchange rate”
[1] Value for the Brent crude oil prices in United States Dollar.
[2] Value for the nominal exchange rate (Brazilian Real per United States Dollar)

Tabela 8: Descriptive statistics

Region	Serie	Mean	Median	Max	Min	Std_Dev	Skewness	Kurtosis
24	Hydrous Ethanol	2.54	2.29	5.64	1.20	0.86	0.98	3.98
24	Diesel Oil	2.52	2.10	4.90	1.31	0.81	0.87	2.89
24	Regular Gasoline	3.21	2.76	6.36	1.94	0.91	1.12	3.60
25	Hydrous Ethanol	2.37	2.21	5.23	1.21	0.73	1.10	4.38
25	Diesel Oil	2.47	2.07	4.62	1.29	0.73	0.80	2.76
25	Regular Gasoline	3.06	2.63	5.92	1.96	0.82	1.12	3.51
26	Hydrous Ethanol	2.49	2.25	5.22	1.43	0.77	0.96	3.49
26	Diesel Oil	2.51	2.10	4.63	1.35	0.75	0.70	2.37
26	Regular Gasoline	3.18	2.74	6.02	2.04	0.83	0.98	3.00
27	Hydrous Ethanol	2.59	2.34	5.40	1.35	0.79	0.98	3.76
27	Diesel Oil	2.49	2.10	4.90	1.36	0.75	0.77	2.63
27	Regular Gasoline	3.33	2.84	6.37	2.06	0.89	1.04	3.34
28	Hydrous Ethanol	2.68	2.48	5.58	1.42	0.78	1.02	3.80
28	Diesel Oil	2.57	2.16	4.84	1.34	0.79	0.75	2.63
28	Regular Gasoline	3.52	2.99	6.58	2.12	0.92	1.03	3.26
29	Hydrous Ethanol	2.39	2.20	5.27	1.16	0.77	1.02	4.06
29	Diesel Oil	2.45	2.09	4.60	1.35	0.70	0.80	2.88
29	Regular Gasoline	3.15	2.73	5.91	1.92	0.81	1.12	3.61
30	Hydrous Ethanol	2.69	2.44	5.90	1.30	0.90	0.92	3.25
30	Diesel Oil	2.61	2.13	5.29	1.36	0.88	0.98	3.03
30	Regular Gasoline	3.44	2.91	6.57	2.08	0.96	1.06	3.18
31	Hydrous Ethanol	2.56	2.33	5.42	1.17	0.81	0.86	3.39
31	Diesel Oil	2.53	2.08	4.95	1.36	0.80	0.90	2.91
31	Regular Gasoline	3.31	2.83	6.19	2.08	0.88	1.16	3.55
32	Hydrous Ethanol	2.51	2.32	5.46	1.10	0.82	0.86	3.44
32	Diesel Oil	2.53	2.10	5.03	1.35	0.80	0.92	3.01
32	Regular Gasoline	3.28	2.81	6.18	2.10	0.85	1.17	3.59
33	Hydrous Ethanol	2.54	2.33	4.60	1.31	0.72	0.77	2.96
33	Diesel Oil	2.51	2.15	4.42	1.33	0.72	0.70	2.49
33	Regular Gasoline	3.19	2.83	5.50	1.99	0.77	1.05	3.40
34	Hydrous Ethanol	2.56	2.32	5.53	1.26	0.76	0.96	3.85
34	Diesel Oil	2.53	2.13	5.07	1.31	0.77	0.87	2.99
34	Regular Gasoline	3.21	2.80	6.09	1.94	0.86	1.09	3.43
35	Hydrous Ethanol	2.46	2.31	4.83	1.33	0.65	0.84	3.10
35	Diesel Oil	2.60	2.19	4.85	1.42	0.76	0.75	2.52
35	Regular Gasoline	3.39	2.98	6.09	2.22	0.79	1.05	3.21

Source: elaborated by the author
Weekly data from May 9th, 2004 to September 4th, 2021
All values are in Brazilian Reais per liter of fuel. Except for the “Brent crude oil” and “Exchange rate”
[1] Value for the Brent crude oil prices in United States Dollar.
[2] Value for the nominal exchange rate (Brazilian Real per United States Dollar)

Tabela 9: Descriptive statistics

Region	Serie	Mean	Median	Max	Min	Std_Dev	Skewness	Kurtosis
36	Hydrous Ethanol	2.59	2.26	5.59	1.38	0.79	1.15	4.03
36	Diesel Oil	2.56	2.17	4.79	1.33	0.76	0.82	2.71
36	Regular Gasoline	3.43	2.93	6.58	2.05	0.90	1.08	3.45
37	Hydrous Ethanol	2.41	2.15	4.92	1.41	0.74	1.09	3.77
37	Diesel Oil	2.44	2.06	4.78	1.29	0.74	0.82	2.80
37	Regular Gasoline	3.25	2.80	6.04	2.02	0.85	1.09	3.37
38	Hydrous Ethanol	2.35	2.15	3.73	1.27	0.63	0.48	1.76
38	Diesel Oil	2.38	2.02	3.92	1.26	0.66	0.51	1.88
38	Regular Gasoline	3.18	2.79	4.89	1.91	0.74	0.64	2.02
39	Hydrous Ethanol	2.42	2.16	4.97	1.27	0.76	0.99	3.62
39	Diesel Oil	2.47	2.10	4.64	1.31	0.72	0.77	2.74
39	Regular Gasoline	3.24	2.81	5.96	1.92	0.85	1.05	3.42
40	Hydrous Ethanol	2.44	2.18	5.23	1.41	0.75	1.09	3.89
40	Diesel Oil	2.52	2.11	4.78	1.34	0.77	0.78	2.65
40	Regular Gasoline	3.33	2.83	6.54	2.07	0.91	1.17	3.63
41	Hydrous Ethanol	2.54	2.24	5.29	1.42	0.80	0.96	3.33
41	Diesel Oil	2.58	2.14	4.90	1.41	0.77	0.81	2.67
41	Regular Gasoline	3.38	2.87	6.41	2.11	0.90	1.10	3.36
42	Hydrous Ethanol	2.35	2.21	4.84	1.25	0.62	0.95	3.98
42	Diesel Oil	2.58	2.21	4.96	1.37	0.79	0.69	2.41
42	Regular Gasoline	3.39	2.96	6.59	2.02	0.97	0.93	2.90
43	Hydrous Ethanol	2.32	2.21	4.62	1.39	0.59	1.12	4.51
43	Diesel Oil	2.55	2.20	4.67	1.36	0.76	0.71	2.44
43	Regular Gasoline	3.37	2.98	6.28	2.18	0.90	1.07	3.21
44	Hydrous Ethanol	2.29	2.14	4.72	1.17	0.66	0.98	3.99
44	Diesel Oil	2.54	2.16	4.74	1.35	0.78	0.76	2.58
44	Regular Gasoline	3.32	2.93	6.35	1.99	0.95	1.01	3.13
45	Hydrous Ethanol	2.34	2.17	4.81	1.23	0.62	0.92	3.63
45	Diesel Oil	2.51	2.14	4.71	1.34	0.76	0.71	2.52
45	Regular Gasoline	3.30	2.94	6.27	2.01	0.93	0.92	2.84
46	Hydrous Ethanol	2.29	2.19	4.69	1.15	0.65	0.95	3.88
46	Diesel Oil	2.52	2.19	4.65	1.34	0.77	0.70	2.45
46	Regular Gasoline	3.22	2.87	6.19	1.89	0.95	0.97	3.02
47	Hydrous Ethanol	2.35	2.18	4.81	1.32	0.66	1.11	4.31
47	Diesel Oil	2.51	2.16	4.68	1.34	0.76	0.75	2.60
47	Regular Gasoline	3.30	2.87	6.33	2.06	0.96	1.04	3.11

Source: elaborated by the author

Weekly data from May 9th, 2004 to September 4th, 2021

All values are in Brazilian Reais per liter of fuel. Except for the “Brent crude oil” and “Exchange rate”

[1] Value for the Brent crude oil prices in United States Dollar.

[2] Value for the nominal exchange rate (Brazilian Real per United States Dollar)

Tabela 10: Descriptive statistics

Region	Serie	Mean	Median	Max	Min	Std_Dev	Skewness	Kurtosis
48	Hydrous Ethanol	2.33	2.18	4.78	1.22	0.67	0.96	3.85
48	Diesel Oil	2.52	2.16	4.77	1.35	0.76	0.76	2.69
48	Regular Gasoline	3.29	2.88	6.27	1.99	0.96	0.98	2.97
49	Hydrous Ethanol	2.28	2.17	4.79	1.12	0.67	1.01	4.29
49	Diesel Oil	2.54	2.15	4.80	1.36	0.78	0.77	2.62
49	Regular Gasoline	3.31	2.89	6.41	1.96	0.97	0.99	3.11
50	Hydrous Ethanol	2.37	2.23	4.79	1.21	0.65	0.95	3.90
50	Diesel Oil	2.55	2.21	4.84	1.41	0.76	0.74	2.58
50	Regular Gasoline	3.32	2.90	6.30	2.02	0.93	0.98	3.03
51	Hydrous Ethanol	2.39	2.26	4.87	1.19	0.68	1.01	4.27
51	Diesel Oil	2.52	2.14	4.76	1.35	0.77	0.75	2.61
51	Regular Gasoline	3.32	2.91	6.40	1.96	0.96	1.01	3.16
52	Hydrous Ethanol	2.42	2.26	5.03	1.33	0.67	1.24	4.92
52	Diesel Oil	2.58	2.23	4.73	1.39	0.77	0.73	2.53
52	Regular Gasoline	3.36	2.96	6.55	2.08	0.96	1.11	3.49
53	Hydrous Ethanol	2.68	2.51	5.43	1.10	0.90	0.63	2.88
53	Diesel Oil	2.56	2.16	4.85	1.45	0.74	0.86	2.94
53	Regular Gasoline	3.32	2.98	6.12	1.99	0.88	1.05	3.43
54	Hydrous Ethanol	2.63	2.46	5.16	1.16	0.87	0.56	2.58
54	Diesel Oil	2.49	2.14	4.41	1.40	0.69	0.70	2.40
54	Regular Gasoline	3.29	2.93	6.10	2.01	0.85	1.05	3.41
55	Hydrous Ethanol	2.58	2.49	5.30	1.03	0.83	0.62	3.13
55	Diesel Oil	2.52	2.18	4.65	1.40	0.70	0.76	2.74
55	Regular Gasoline	3.22	2.88	6.18	1.97	0.83	1.17	3.96
56	Hydrous Ethanol	2.66	2.52	5.59	1.23	0.82	0.74	3.33
56	Diesel Oil	2.57	2.20	4.79	1.47	0.70	0.77	2.64
56	Regular Gasoline	3.37	2.99	6.49	2.10	0.86	1.14	3.82
57	Hydrous Ethanol	2.48	2.21	5.65	1.09	0.90	0.92	3.38
57	Diesel Oil	2.46	2.07	4.59	1.28	0.74	0.69	2.45
57	Regular Gasoline	3.43	2.93	6.78	2.08	1.00	1.03	3.12
58	Hydrous Ethanol	2.57	2.27	5.73	1.10	0.94	0.77	2.82
58	Diesel Oil	2.52	2.11	4.69	1.34	0.75	0.67	2.40
58	Regular Gasoline	3.45	3.03	6.58	2.05	1.00	0.94	2.94
59	Hydrous Ethanol	2.56	2.32	5.50	1.10	0.89	0.81	2.93
59	Diesel Oil	2.49	2.10	4.55	1.32	0.74	0.67	2.47
59	Regular Gasoline	3.41	2.94	6.72	2.05	0.98	1.05	3.24

Source: elaborated by the author
Weekly data from May 9th, 2004 to September 4th, 2021
All values are in Brazilian Reais per liter of fuel. Except for the “Brent crude oil” and “Exchange rate”

[1] Value for the Brent crude oil prices in United States Dollar.

[2] Value for the nominal exchange rate (Brazilian Real per United States Dollar)

Tabela 11: Descriptive statistics

Region	Serie	Mean	Median	Max	Min	Std_Dev	Skewness	Kurtosis
60	Hydrous Ethanol	2.64	2.38	5.61	1.19	0.93	0.85	3.16
60	Diesel Oil	2.59	2.17	4.75	1.39	0.78	0.69	2.42
60	Regular Gasoline	3.50	3.04	6.69	2.12	1.01	1.00	3.13
61	Hydrous Ethanol	2.68	2.40	5.84	1.15	0.99	0.85	3.12
61	Diesel Oil	2.59	2.16	4.80	1.42	0.79	0.74	2.43
61	Regular Gasoline	3.52	3.03	6.71	2.13	1.02	0.98	3.07
62	Hydrous Ethanol	2.54	2.27	5.44	1.04	0.91	0.84	3.12
62	Diesel Oil	2.52	2.12	4.57	1.36	0.75	0.69	2.39
62	Regular Gasoline	3.36	2.92	6.49	1.98	1.00	0.99	3.10
63	Hydrous Ethanol	1.92	1.84	4.38	0.71	0.68	0.82	3.69
63	Diesel Oil	2.46	2.10	4.62	1.35	0.71	0.84	2.89
63	Regular Gasoline	3.12	2.75	5.83	1.89	0.81	1.02	3.34
64	Hydrous Ethanol	1.95	1.85	4.47	0.77	0.69	0.82	3.62
64	Diesel Oil	2.47	2.09	4.63	1.33	0.72	0.81	2.84
64	Regular Gasoline	3.12	2.72	5.87	1.92	0.83	0.99	3.21
65	Hydrous Ethanol	1.95	1.82	4.30	0.76	0.67	0.86	3.69
65	Diesel Oil	2.50	2.13	4.63	1.38	0.72	0.78	2.74
65	Regular Gasoline	3.17	2.84	5.82	1.93	0.82	1.00	3.35
66	Hydrous Ethanol	1.98	1.86	4.42	0.77	0.69	0.84	3.61
66	Diesel Oil	2.50	2.12	4.63	1.37	0.72	0.79	2.76
66	Regular Gasoline	3.11	2.74	5.75	1.92	0.80	1.01	3.28
67	Hydrous Ethanol	1.94	1.81	4.45	0.79	0.69	0.94	3.87
67	Diesel Oil	2.49	2.13	4.55	1.37	0.70	0.80	2.83
67	Regular Gasoline	3.11	2.69	5.80	1.94	0.81	1.06	3.33
68	Hydrous Ethanol	1.93	1.84	4.38	0.72	0.69	0.81	3.54
68	Diesel Oil	2.44	2.08	4.52	1.35	0.69	0.84	2.93
68	Regular Gasoline	3.05	2.70	5.69	1.87	0.80	1.05	3.43
69	Hydrous Ethanol	1.95	1.85	4.40	0.77	0.69	0.83	3.55
69	Diesel Oil	2.45	2.09	4.53	1.35	0.70	0.82	2.88
69	Regular Gasoline	3.04	2.67	5.67	1.88	0.78	1.05	3.40
70	Hydrous Ethanol	1.95	1.84	4.33	0.81	0.65	0.89	3.84
70	Diesel Oil	2.50	2.13	4.66	1.36	0.71	0.77	2.76
70	Regular Gasoline	3.13	2.76	5.79	1.93	0.81	1.04	3.33
71	Hydrous Ethanol	1.99	1.89	4.30	0.81	0.68	0.81	3.58
71	Diesel Oil	2.50	2.12	4.70	1.38	0.73	0.78	2.63
71	Regular Gasoline	3.13	2.76	5.62	1.92	0.80	0.99	3.15

Source: elaborated by the author
Weekly data from May 9th, 2004 to September 4th, 2021
All values are in Brazilian Reais per liter of fuel. Except for the "Brent crude oil" and "Exchange rate"
[1] Value for the Brent crude oil prices in United States Dollar.
[2] Value for the nominal exchange rate (Brazilian Real per United States Dollar)

Tabela 12: Descriptive statistics

Region	Serie	Mean	Median	Max	Min	Std_Dev	Skewness	Kurtosis
72	Hydrous Ethanol	1.96	1.84	4.44	0.77	0.68	0.90	3.70
72	Diesel Oil	2.50	2.12	4.69	1.36	0.73	0.81	2.76
72	Regular Gasoline	3.12	2.70	5.78	1.92	0.81	1.06	3.28
73	Hydrous Ethanol	1.97	1.87	4.38	0.76	0.70	0.80	3.46
73	Diesel Oil	2.47	2.10	4.60	1.35	0.71	0.83	2.86
73	Regular Gasoline	3.03	2.70	5.64	1.85	0.80	1.02	3.31
74	Hydrous Ethanol	2.03	1.91	4.48	0.79	0.71	0.78	3.40
74	Diesel Oil	2.47	2.12	4.65	1.38	0.70	0.88	2.99
74	Regular Gasoline	3.06	2.70	5.69	1.93	0.78	1.08	3.47
75	Hydrous Ethanol	1.99	1.89	4.44	0.74	0.70	0.78	3.46
75	Diesel Oil	2.48	2.12	4.59	1.39	0.70	0.82	2.86
75	Regular Gasoline	3.05	2.70	5.71	1.84	0.81	1.03	3.35
76	Hydrous Ethanol	2.04	1.89	4.60	0.84	0.71	0.80	3.33
76	Diesel Oil	2.50	2.10	4.72	1.38	0.73	0.85	2.78
76	Regular Gasoline	3.11	2.69	5.89	1.91	0.80	1.02	3.23
99	Hydrous Ethanol	2.38	2.16	4.58	1.19	0.72	0.83	2.78
99	Diesel Oil	2.66	2.28	4.54	1.51	0.70	0.63	2.26
99	Regular Gasoline	3.25	2.86	5.89	2.03	0.76	1.28	4.19
100	Hydrous Ethanol	2.50	2.34	4.87	1.25	0.75	0.81	3.00
100	Diesel Oil	2.73	2.36	4.87	1.51	0.72	0.58	2.23
100	Regular Gasoline	3.38	3.00	6.26	2.14	0.79	1.13	3.85
101	Hydrous Ethanol	2.52	2.30	5.17	1.30	0.75	0.95	3.40
101	Diesel Oil	2.76	2.36	4.92	1.54	0.73	0.70	2.62
101	Regular Gasoline	3.36	2.95	6.23	2.14	0.81	1.25	4.32
102	Hydrous Ethanol	2.31	2.14	4.61	1.50	0.60	1.14	4.38
102	Diesel Oil	2.92	2.53	5.26	1.67	0.79	0.75	2.52
102	Regular Gasoline	3.54	3.14	6.21	2.45	0.78	1.21	3.62
103	Hydrous Ethanol	2.10	1.98	4.60	1.20	0.58	1.14	4.97
103	Diesel Oil	2.74	2.40	4.94	1.55	0.73	0.74	2.63
103	Regular Gasoline	3.34	2.96	6.16	2.28	0.76	1.29	4.04
104	Hydrous Ethanol	2.16	1.98	4.62	1.35	0.61	1.24	4.85
104	Diesel Oil	2.75	2.38	4.99	1.55	0.77	0.77	2.61
104	Regular Gasoline	3.37	2.98	6.10	2.34	0.74	1.34	4.08
105	Hydrous Ethanol	2.20	2.02	4.72	1.12	0.71	1.12	4.25
105	Diesel Oil	2.55	2.12	4.84	1.42	0.78	0.82	2.70
105	Regular Gasoline	3.31	2.98	6.08	2.03	0.87	1.10	3.51

Source: elaborated by the author

Weekly data from May 9th, 2004 to September 4th, 2021

All values are in Brazilian Reais per liter of fuel. Except for the “Brent crude oil” and “Exchange rate”

[1] Value for the Brent crude oil prices in United States Dollar.

[2] Value for the nominal exchange rate (Brazilian Real per United States Dollar)

Tabela 13: Descriptive statistics

Region	Serie	Mean	Median	Max	Min	Std_Dev	Skewness	Kurtosis
106	Hydrous Ethanol	2.12	1.95	4.66	1.02	0.72	0.94	3.71
106	Diesel Oil	2.49	2.11	4.76	1.39	0.76	0.79	2.67
106	Regular Gasoline	3.23	2.82	6.33	1.93	0.93	1.06	3.43
107	Hydrous Ethanol	2.32	2.15	4.86	1.28	0.66	1.24	4.89
107	Diesel Oil	2.58	2.18	4.85	1.45	0.78	0.87	2.87
107	Regular Gasoline	3.29	2.87	6.48	2.02	0.90	1.20	4.01
108	Hydrous Ethanol	2.22	2.03	4.68	1.16	0.71	1.01	3.87
108	Diesel Oil	2.57	2.19	4.80	1.45	0.77	0.81	2.70
108	Regular Gasoline	3.33	2.94	6.24	2.05	0.90	1.09	3.49
109	Hydrous Ethanol	2.49	2.27	5.41	1.22	0.76	0.92	3.75
109	Diesel Oil	2.58	2.13	4.73	1.44	0.78	0.68	2.32
109	Regular Gasoline	3.24	2.88	6.48	1.94	0.84	1.14	4.08
77	Hydrous Ethanol	2.11	1.95	4.65	0.86	0.74	0.77	3.21
77	Diesel Oil	2.45	2.10	4.45	1.37	0.67	0.77	2.74
77	Regular Gasoline	3.21	2.80	5.90	1.94	0.85	0.93	2.99
78	Hydrous Ethanol	2.05	1.90	4.57	0.82	0.72	0.80	3.34
78	Diesel Oil	2.42	2.09	4.41	1.35	0.67	0.80	2.91
78	Regular Gasoline	3.11	2.74	5.69	1.90	0.82	0.92	3.03
79	Hydrous Ethanol	2.13	1.95	4.84	0.87	0.72	1.00	4.10
79	Diesel Oil	2.45	2.10	4.46	1.36	0.68	0.80	2.85
79	Regular Gasoline	3.15	2.78	5.98	2.00	0.85	1.04	3.42
80	Hydrous Ethanol	2.17	2.02	4.76	0.89	0.73	0.72	3.06
80	Diesel Oil	2.43	2.10	4.31	1.37	0.65	0.64	2.42
80	Regular Gasoline	3.22	2.85	5.87	1.98	0.83	0.82	2.70
81	Hydrous Ethanol	2.23	2.06	4.66	1.00	0.71	0.87	3.44
81	Diesel Oil	2.43	2.09	4.43	1.35	0.66	0.81	2.90
81	Regular Gasoline	3.18	2.80	5.78	1.98	0.80	0.97	3.08
82	Hydrous Ethanol	2.14	2.02	4.62	0.93	0.72	0.74	3.11
82	Diesel Oil	2.47	2.13	4.48	1.40	0.68	0.77	2.73
82	Regular Gasoline	3.20	2.83	5.85	1.98	0.84	0.90	2.87
83	Hydrous Ethanol	2.23	2.05	4.69	0.99	0.72	0.79	3.16
83	Diesel Oil	2.45	2.12	4.32	1.38	0.65	0.71	2.50
83	Regular Gasoline	3.21	2.87	5.87	1.99	0.82	0.94	3.00
84	Hydrous Ethanol	2.14	2.01	4.68	0.92	0.70	0.87	3.56
84	Diesel Oil	2.40	2.06	4.35	1.36	0.65	0.82	2.94
84	Regular Gasoline	3.07	2.74	5.69	1.90	0.79	1.02	3.36

Source: elaborated by the author
Weekly data from May 9th, 2004 to September 4th, 2021
All values are in Brazilian Reais per liter of fuel. Except for the "Brent crude oil" and "Exchange rate"
[1] Value for the Brent crude oil prices in United States Dollar.
[2] Value for the nominal exchange rate (Brazilian Real per United States Dollar)

Tabela 14: Descriptive statistics

Region	Series	Mean	Median	Max	Min	Std_Error	Skewness	Kurtosis
85	Hydrous Ethanol	2.23	2.09	4.75	1.00	0.71	0.88	3.48
85	Diesel Oil	2.46	2.13	4.48	1.38	0.66	0.78	2.77
85	Regular Gasoline	3.19	2.82	5.87	1.99	0.79	0.95	3.14
86	Hydrous Ethanol	2.61	2.45	5.45	1.25	0.85	0.74	3.06
86	Diesel Oil	2.52	2.17	4.63	1.39	0.68	0.76	2.76
86	Regular Gasoline	3.20	2.87	5.95	2.04	0.78	0.98	3.35
87	Hydrous Ethanol	2.52	2.40	5.18	1.21	0.81	0.72	3.06
87	Diesel Oil	2.51	2.19	4.77	1.42	0.69	0.80	2.91
87	Regular Gasoline	3.11	2.79	5.62	2.05	0.71	1.06	3.45
88	Hydrous Ethanol	2.59	2.45	5.40	1.28	0.83	0.80	3.30
88	Diesel Oil	2.52	2.17	4.59	1.43	0.66	0.75	2.67
88	Regular Gasoline	3.17	2.79	5.87	2.09	0.77	1.15	3.75
89	Hydrous Ethanol	2.51	2.39	5.22	1.18	0.82	0.81	3.31
89	Diesel Oil	2.50	2.15	4.54	1.40	0.67	0.78	2.88
89	Regular Gasoline	3.10	2.75	5.70	2.02	0.74	1.27	4.23
90	Hydrous Ethanol	2.51	2.36	5.26	1.28	0.84	0.77	3.06
90	Diesel Oil	2.52	2.21	4.57	1.38	0.70	0.72	2.73
90	Regular Gasoline	3.13	2.73	5.88	2.14	0.78	1.15	3.75
91	Hydrous Ethanol	2.55	2.45	5.21	1.17	0.81	0.69	3.01
91	Diesel Oil	2.45	2.14	4.42	1.40	0.65	0.86	3.07
91	Regular Gasoline	3.13	2.79	5.73	2.06	0.73	1.19	3.93
92	Hydrous Ethanol	2.80	2.47	5.95	1.25	1.02	0.85	2.89
92	Diesel Oil	2.55	2.21	4.59	1.42	0.67	0.80	2.75
92	Regular Gasoline	3.35	2.87	6.28	2.16	0.89	1.08	3.25
93	Hydrous Ethanol	2.82	2.51	6.11	1.31	1.01	0.92	3.21
93	Diesel Oil	2.57	2.23	4.59	1.44	0.67	0.81	2.84
93	Regular Gasoline	3.35	2.91	6.28	2.22	0.89	1.12	3.44
94	Hydrous Ethanol	2.76	2.43	5.98	1.13	1.05	0.81	2.76
94	Diesel Oil	2.57	2.21	4.79	1.40	0.72	0.88	3.02
94	Regular Gasoline	3.33	2.90	6.41	2.17	0.91	1.18	3.59
95	Hydrous Ethanol	2.78	2.44	5.99	1.19	1.03	0.85	2.96
95	Diesel Oil	2.55	2.21	4.60	1.40	0.68	0.82	2.93
95	Regular Gasoline	3.30	2.87	6.16	2.18	0.88	1.13	3.37
96	Hydrous Ethanol	2.70	2.40	5.91	1.09	0.99	0.88	3.08
96	Diesel Oil	2.48	2.17	4.41	1.38	0.65	0.83	2.88
96	Regular Gasoline	3.21	2.79	6.22	2.01	0.88	1.16	3.55

Source: elaborated by the author

Weekly data from May 9th, 2004 to September 4th, 2021

All values are in Brazilian Reais per liter of fuel. Except for the “Brent crude oil” and “Exchange rate”

[1] Value for the Brent crude oil prices in United States Dollar.

[2] Value for the nominal exchange rate (Brazilian Real per United States Dollar)

Tabela 15: Descriptive statistics

Region	Serie	Mean	Median	Max	Min	Std_Dev	Skewness	Kurtosis
97	Hydrous Ethanol	2.93	2.57	6.41	1.37	1.08	0.95	3.22
97	Diesel Oil	2.62	2.24	4.90	1.44	0.75	0.91	3.01
97	Regular Gasoline	3.50	2.93	6.63	2.23	0.98	1.15	3.51
98	Hydrous Ethanol	2.87	2.57	6.20	1.35	1.01	0.92	3.19
98	Diesel Oil	2.58	2.23	4.70	1.41	0.69	0.83	2.91
98	Regular Gasoline	3.44	2.94	6.44	2.25	0.89	1.10	3.39
	Brent crude oil	73.61	67.75	143.05	21.61	25.23	0.49	2.19
	Exchange rate BRL/USD	2.78	2.34	5.85	1.55	1.06	1.04	3.28

Source: elaborated by the author

Weekly data from May 9th, 2004 to September 4th, 2021

All values are in Brazilian Reais per liter of fuel. Except for the "Brent crude oil" and "Exchange rate"

[1] Value for the Brent crude oil prices in United States Dollar.

[2] Value for the nominal exchange rate (Brazilian Real per United States Dollar)

Tabela 16: Grouping of the regions

Region[1]	Group[2]	Region[1]	Group[2]	Region[1]	Group[2]	Region[1]	Group[2]
1504	1598	3111	3111	2104	2104	3303	3303
2605	2699	3201	3201	4101	4101	2204	2204
2102	2102	2201	2201	1506	1597	2501	2599
3508	3508	2901	2901	4104	4104	2602	2602
2601	2601	3502	3502	5001	5099	2101	2101
5204	5204	3515	3515	2304	2304	5202	5202
2904	2904	2504	2504	2307	2307	1202	1299
4306	4306	2302	2398	4301	4301	5105	5105
1501	151699	3306	3306	1201	1299	4303	4303
3110	3110	3107	3107	2701	2701	3104	3199
4110	4110	2603	2699	3103	3199	4109	4199
5101	5101	4302	4302	1702	1702	1502	151699
1505	1597	4205	4205	3109	3109	1402	1499
4305	4305	1401	1499	5004	5099	1601	151699
3507	3507	3106	3106	2907	2907		
1503	1598	5301	5301	4107	4107		
5203	5203	2906	2906	3509	3509		
3305	3305	4201	4201	2604	2699		
3511	3599	3513	3513	1101	1101		
4103	4103	4307	4307	4108	4199		
2803	2803	5104	5104	1304	1399		
3503	3503	4304	4304	2306	2399		
3202	3202	3204	3204	2301	2301		
1701	1701	2402	2402	2802	2802		
3105	3105	5205	5205	1303	1399		
2702	2702	2905	2905	3514	3599		
4206	4206	2503	2599	3301	3301		
3505	3505	2903	2903	3102	3102		
3506	3506	5002	5002	4202	4202		
3304	3304	2202	2202	3101	3101		
1102	1102	4102	4102	2902	2902		
3510	3510	3302	3302	4203	4203		
4106	4106	3108	3108	2305	2399		
3512	3512	3203	3203	1602	151699		
3504	3504	4105	4105	2703	2703		
2103	2199	3112	3112	2401	2401		
4204	4204	3501	3501	2404	2404		
2105	2199	2303	2398	5003	5003		

Source: elaborated by the author
[1] Mesoregion Code according to Brazilian Institute of Geography and Statistics
[2] Group code.

Tabela 17: P-values for the Diebold-Mariano test for predictive accuracy

	M8	**M6**	**M16**	**M14**	**M7**	**M15**	**M1**	**M9**	**VECM2**	**VECM**	**M12**	**M4**	**M10**	**M2**	**M13**	**M5**	**M11**	**M3**
M8		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
M6	0.000		1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
M16	0.000	0.000		0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
M14	0.000	0.006	1.000		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
M7	0.000	0.000	0.000	0.000		0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
M15	0.000	0.000	0.000	0.000	0.006		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
M1	0.000	0.000	0.000	0.000	0.000	0.000		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
M9	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
VECM2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
VECM	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
M12	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		1.000	1.000	1.000	1.000	1.000	1.000	1.000
M4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		1.000	1.000	1.000	1.000	1.000	1.000
M10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		1.000	1.000	1.000	1.000	1.000
M2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		1.000	1.000	1.000	1.000
M13	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		1.000	1.000	1.000
M5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		1.000	1.000
M11	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		1.000
M3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	

Diebold-Mariano test with one step ahead forecast
Loss function is the squared-error loss
Alternative hypothesis: row model error is greater than column model error