Integration in Gasoline and Ethanol Markets in Brazil over Time and Space under the Flex-fuel Technology

Hector M. Nuñez and Jesús Otero***

ABSTRACT

We employ a pair-wise approach to analyse regional integration in the gasoline and ethanol markets in Brazil. Using weekly price data for these two fuels at the state level over a period of almost ten years, we find that more than half of the fuel price differentials are stationary, which reveals the importance of allowing for spatial considerations when testing for market integration. We also find that the speed at which prices converge to the long-run equilibrium depends upon the distance between states, the differential in sugarcane mills density between states, and the similarity between tax regimes. Other demand and supply factors such as population density, gas stations density, sugarcane mills density and GDP per capita are not statistically significant.

Keywords: Gasoline, Ethanol, Prices, Market integration, Distance

http://dx.doi.org/10.5547/01956574.38.2.hnun

1. INTRODUCTION

In 1973 Brazil, a country heavily dependent on petroleum imports, was severely hit by the first oil crisis. To mitigate the effects of the crisis, the Brazilian government embarked on two ambitious programmes to substitute imported oil with domestic energy sources. The first programme had the objective to increase petroleum reserves mainly through exploration activities initially in offshore shallow waters and subsequently in offshore deep waters. The second programme, established in November 1975 and known as Programa Nacional do Álcool or Pró-Álcool for short, had the purpose of producing large quantities of ethanol from biomass (e.g. sugarcane, cassava and sorghum) as a substitute for gasoline by providing economic incentives to ethanol producers and consumers. For a variety of reasons, including low international prices for sugar and idle capacity for distillation, sugarcane became the sole source of ethanol.[1]

1. Sugarcane production in Brazil dates back to 1532. It is grown on a six-year cycle, where one planting year is followed by an initial harvest after 12–18 months, and four or five succeeding harvests. Once sugarcane is harvested, the stalks are cut and taken to the mills while the roots and rhizomes are left in the soil for sugarcane regrowth. Transportation of sugarcane to the mill needs to be done within 24 hours of cutting; otherwise there is a loss in the sugar content of the sugarcane. Stalks are cleaned at the mill (either through washing or dry-cleaning) and subsequently passed through a set of knives graders and choppers, followed by the grinder. Then the cane pieces are crushed and soaked with water to separate the juice from the bagasse. Clarification and evaporation then take place, leaving the syrup or raw material ready for ethanol or sugar production. For ethanol, syrup is first converted to fructose, then fermented and finally distilled to produce hydrous ethanol (93% ethyl alcohol concentration), which can also be dehydrated to produce anhydrous ethanol (99.3% concentration) to be blended with gasoline or exported; the interested reader is referred to PECEGE (2009) for more details on ethanol production.

* Corresponding author. Department of Economics. Centro de Investigación y Docencia Económicas, Circuito Tecnopolo Norte 117, Aguascalientes, AGS, 20313, México. E-mail: hector.nunez@cide.edu.

** Facultad de Economía, Universidad del Rosario, Calle 12 C # 4-69, Bogotá, Colombia. E-mail: jesus.otero@urosario.edu.co.

Historical accounts of the ethanol programme in Brazil indicate that Pró-Álcool initially focussed on the production of both anhydrous ethanol and hydrous ethanol; see, for example, Rosillo-Calle and Cortez (1998) and Goldemberg (2006). Anhydrous ethanol has remained compulsory as an additive to gasoline in blends of varying proportions that have been increased over the years from 10% to 27%; the usage of this fuel requires no modifications in the car engines and the blend mandate is still currently in place. On the other hand, the Brazilian automotive industry developed ethanol-dedicated vehicles that use exclusively 100% hydrous ethanol. According to the Brazilian National Association of Motor Vehicle Manufacturers, ethanol car sales had an initial rapid increase during the 1980s, reaching 93.6% of total sales of new cars in 1987.[2] However, Rosillo-Calle and Cortez (1998) and Salvo and Huse (2011) point out several reasons why the relative importance of ethanol-dedicated vehicles sales reduced rapidly throughout the 1990s, such as the elimination of subsidies and price supports, the deregulation of the ethanol industry, low international crude oil prices, high sugar prices in world markets, and oil discoveries off the Brazilian coast. Consequently, ethanol production did not increase during those years.

In March 2003 a major technological change took place in the Brazilian automotive industry with the introduction to the market of the flex-fuel vehicles, which are capable of running on any blend of gasoline and hydrous ethanol. Salvo and Huse (2013) point out that by 2010 flex-fuel vehicles amount to almost a third of the light vehicle stock, and approximately 50% of the light duty vehicle-miles travelled in the country. There are 12 automakers that offer over 90 models of flex fuel vehicles in Brazil that account for 50% of the current fleet. This percentage is expected to grow to 86% by 2020 (Jank, 2011). Pacini and Silveira (2011) indicate that the rapid acceptance of the flex-fuel technology by Brazilian consumers depended on the fact that they were now able to react to price signals and switch from one fuel to another on a daily basis. This is in sharp contrast with the previous situation in which consumers could only take into account price signals when deciding which type of car to purchase, that is gasoline or ethanol driven. Pacini and Silveira (2011) conclude that the introduction of the flex-fuel technology has opened a major connection between the gasoline and ethanol markets in Brazil.

Along with this technological change, both federal and state governments have provided lower tax rates to ethanol relative to gasoline, which have boosted domestic ethanol consumption during the last decade, reaching its peak in 2010 with 24.4 billion litres in contrast to the 22.8 billion litres of gasoline in the same year.[3] Infrastructure to sell ethanol has kept pace with the growth in its production. In 2014 there were 409 mills distributed across 23 states. These mills, most of which are located in the Southeaster region of the country, deliver the ethanol to PETROBRAS and 205 authorized distributors, which in turn send it to the nine ethanol collection centres, where anhydrous ethanol is blended. Then, blended gasoline and hydrous ethanol are delivered to 409 retail centres located across all states, which in turn sell the fuels to the stations (see Valdes (2011) for more details on the ethanol industry and market). About 92% of the fuel stations in Brazil sell both gasoline and hydrous ethanol (ANP, 2013). On the other hand, oil is refined in 16 refineries located in nine states and subsequently transported to the retail centres through a network of (more than 11,000 km of) pipelines and tanker trucks; see Costa and Neves (2015) for more details about oil and gasoline transportation.

It is important to notice that the fuel transportation sector has been quite dynamic in recent years. In response to shortages in the supply of domestic ethanol, the ethanol blending percentage

2. See Anúario da Indústria Automobilística Brasileira, available at http://www.anfavea.com.br/anuario.html.

3. See the Balanço Energético Nacional (BEN) produced by the Empresa de Pesquisa Energética, which is available at https://ben.epe.gov.br/.

has been reduced for some periods and the federal tax rate on gasoline lowered so that it has been practically eliminated. In contrast, when sugarcane production and supply of ethanol are high, the government has increased the blending mandate. As to gasoline, to curb the substantial losses from subsidized domestic gasoline sales by refineries, the past few years have witnessed price increases of more than 10%, after a long period of wholesale price controls. Furthermore, in Brazil the ethanol and sugar markets are strongly related since sugarcane is the common feedstock for ethanol and sugar production. In the past decade, the global sugar market has been quite dynamic too. The strong demand for sugar and associated rising prices observed some years ago have now been replaced by a significant oversupply and declining prices. This has had an important impact on the Brazilian sugar industry, and thus in the ethanol sector as well. Bearing in mind that historically PETROBRAS has controlled gasoline prices from the refinery to the retail centres, as opposed to the more competitive ethanol market, arbitrage opportunities in the fuel market are expected to occur after ethanol is delivered to the collection centres.

This paper aims to further our understanding of the extent of spatial integration in the markets for gasoline and ethanol in Brazil, after the introduction of the flex-fuel car technology. According to Fackler and Goodwin (2001), the extent of spatial market integration has often been examined by investigating the validity of the law of one price, either by testing whether the prices of identical products traded in different locations are the same once they are converted to a common currency (that is, the absolute version of the law), or by testing whether price discrepancies are stationary or mean reverting (that is, the relative version of the law).

The econometric modelling strategy that we propose in this paper offers two distinguishing features with respect to the existing literature. First, we study the interaction between gasoline and ethanol prices at the state level by adapting the Pesaran (2007) pairwise procedure to the study of market integration. The idea underlying the pairwise procedure is that given a sample of N prices, unit root tests are conducted on all $(N(N-1)/2)$ price differentials, so that the total number of stationary differentials can be determined. The use of the pairwise procedure offers the advantage that, by calculating all possible price differentials, it does not involve the choice (in some cases arbitrary) of a base or benchmark price with respect to which all other prices ought to be measured. Bearing in mind that Brazil is a federation that consists of 27 states, a pairwise analysis of 27 gasoline price series and 27 ethanol price series, that is 54 price series, yields a total of 1,431 possible price differentials that can be computed. Consequently, the application of this modelling strategy not only allows us to study the possibility of spatial integration within the gasoline and ethanol markets, but also between them. To the best of our knowledge, this is perhaps the most comprehensive analysis of integration in the Brazilian gasoline and ethanol markets available in the literature.

The second novelty of our modelling strategy is that once the number of stationary price differentials is determined, in a subsequent stage of the analysis we employ information from geographical and economic variables to explain differences in the speed of convergence towards long-run equilibrium across the pairwise price differentials. More specifically, in the event of regional shocks to gasoline or ethanol prices, we investigate whether the speed of adjustment towards long-run equilibrium is fastest between contiguous as opposed to more distant or non-contiguous states, by relying on the distance between states as an explanatory variable. However, we also evaluate the role played by variables such as sugarcane mills density, population density, real per-capita GDP, gas stations density, and gasoline and ethanol tax regimes at the state level in influencing the speed of adjustment of price differentials.

The paper proceeds as follows. Section 2 presents a brief review of the literature. Section 3 outlines the econometric modelling strategy. Section 4 describes the data and presents the results of the empirical analysis. Section 5 offers concluding remarks.

2. A BRIEF REVIEW OF THE LITERATURE

Numerous studies have addressed the existence of integration in gasoline markets including, inter alia, Stigler and Sherwin (1985), Spiller and Huang (1986), Paul, Miljkovic, and Ipe (2001) and Holmes, Otero, and Panagiotidis (2013) for the United States, and Suvankulov, Lau, and Ogucu (2012) for Canada.[4]

In spite of the extensive literature, the above studies leave scope for further research and analysis. Indeed, the study of the Brazilian experience with gasoline and ethanol is interesting because it offers the possibility of investigating two highly connected markets. As indicated in the previous section, prior to the introduction of the flex-fuel car technology, adapting car engines from gasoline-fuel to alcohol-fuel was expensive and therefore substitution between these fuels was limited. This fact is confirmed when one examines the values of the long-run own-price elasticity of demand for gasoline that have been estimated for Brazil in several studies, where low (high) values indicate low (high) substitutability. Despite the methodological differences that exist among studies, it is possible to identify that as the data that are used in the empirical studies move closer to our sample period, the estimates of the elasticity that are reported increase in absolute value. For instance, for the period before the flex-fuel era, own-price elasticity of demand estimates for gasoline include -0.2 by Assis and Rodrigues Lopes (1980), who use a pooled estimation for five regions in Brazil from 1970 to 1977. Gately and Streifel (1997), using data from 1971 to 1993 for thirty-seven developing countries, find that in Brazil the responses of the demand to the maximum historical price, price cuts and prices recoveries are respectively -0.19, -0.38 and -0.49. Rogat and Sterner (1998) find a value of the elasticity of demand of -0.98 for Brazil using a sample of Latin American countries for the period 1960-1994. Burnquist and Bacchi (2002) estimate a value of -0.31 using a cointegration approach with annual data from 1973 to 1998. Alves and Bueno (2003) obtain a value for the own-price elasticity equal to -0.46 with annual data from 1974 to 1999.[5] Roppa (2005) finds a long-run elasticity of -0.63 using also a cointegration technique for twenty-two annual observations between 1979 and 2000.

Studies which provide estimates of the own-price elasticity of demand for gasoline in the flex-fuel era (i.e. after 2003) include Schünemann (2007), who finds a long-run elasticity of -0.29 using cointegration procedures with annual data from 1970 to 2005. Silva, Tiryaki, and Pontes (2009) estimate a panel data model with fixed effects for twenty-seven Brazilian states using quarterly data 2001 to 2006, and find a long-run price elasticity of -0.57. Much larger values (in absolute value) are obtained by Santos (2013) and de Freitas and Kaneko (2011a) with estimates of -1.2 and -1.8, respectively. The results of the former are based on a panel cointegration approach using quarterly data from July 2001 to December 2010, while those of the latter are based on applying a time-series cointegration analysis to monthly data from January 2003 to June 2010. The higher estimated values of the cross-price elasticity of gasoline and ethanol demand obtained by de Freitas and Kaneko (2011a) and Santos (2013) provide support for the view that the flex-fuel car technology, by eliminating switching costs, has indeed increased consumer choices and stimulated competition between fuels. Introducing a regional perspective in the ethanol demand analysis, de Freitas and Kaneko (2011b) divide the country in two regions, namely the centre-south (CS) and the north-northeast (NN). The distinguishing feature between the two regions is that the economic and social

4. A different strand of the literature has studied the transmission of shocks between wholesale and retail prices. For example, see da Silva et al. (2014) for a recent analysis using prices of regular gasoline at the gas stations and at the distributors in 134 municipalities throughout Brazil.

5. This coefficient, however, is statistically different from zero at the 10% level based on a one-sided *t*-test.

development indicators in the CS region are higher than in the NN region. The results show that ethanol demand in the CS region is characterised by higher price elasticities compared to the lower values that are obtained for the NN region.

Ferreira, de Almeida Prado, and da Silveira (2009) and Salvo and Huse (2011) build theoretical models to accommodate the presence of the flex-fuel technology in the Brazilian fuel market. These models indicate that the possibility of substituting gasoline and ethanol as alternative fuels implies the existence of a long-run cointegration relationship between the prices of these two fuels. Moreover, given the lower energy content (or fuel economy) of ethanol relative to blended gasoline, pricing parity occurs when the price of ethanol is approximately equal to 70% of the price of gasoline.[6]

Empirical support for the existence of cointegration between gasoline and ethanol prices is somewhat mixed. Using data for the country as a whole, Ferreira, de Almeida Prado, and da Silveira (2009) and Du and Carriquiry (2013) find that the differential between the ethanol and gasoline price series is stationary. Ferreira, de Almeida Prado, and da Silveira (2009) further report that causality (in the Granger sense) between gasoline and ethanol runs from the former to the later, but not vice versa. Tello-Gamarra (2009) carries out cointegration tests between ethanol and gasoline markets in Brazil using monthly prices from 2003 to 2008. This author finds evidence of cointegration in the long-run, causality in both directions, and that the intensity of price transmission from gasoline to ethanol is 1 to 2.74. In turn, Serra, Zilberman, and Gil (2011), employing data on weekly international crude oil prices, and Brazilian ethanol and sugar prices, find that in the long run an increase in crude oil prices leads the system to a new equilibrium characterised by higher ethanol prices. Moreover, ethanol prices respond to departures from the long-run equilibrium relationship, while crude oil and sugar prices do not. Barros, Gil-Alana, and Wanke (2014) use fractional integration techniques to study the degree of persistence in the ethanol to gasoline price ratio, and find that it is not stationary.

Within a regional framework, Salvo and Huse (2011) test for the existence of cointegration between gasoline and ethanol prices using state-level price data, and find that support for the hypothesis occurs in seven out of 27 states. In six additional cases the gasoline and ethanol price series are found to be stationary in levels, so that the difference between the two is, by definition, also a stationary series. Although Salvo and Huse (2011) recognise in their theoretical model that distance may play a role as a factor that determines the speed at which gasoline and ethanol prices adjust to their pricing parity, spatial considerations are not present in their empirical analysis. Indeed, they only examine the time-series properties of $(p_i^g - p_i^e)$, where p_i^g and p_i^e denote the prices of gasoline and ethanol in state i, respectively. However, if one allows for spatial considerations, then it is necessary to examine the time-series properties not only of $(p_i^g - p_i^e)$ but also of $(p_i^g - p_j^e)$, $(p_i^e - p_j^e)$ and $(p_i^g - p_j^g)$. Overall, for every pair of states there will be a total of six price differentials worth considering when analysing arbitrage opportunities. For instance, for Distrito Federal (DF) and Goiás (GO) the corresponding price differentials are $(p_{\mathrm{GO}}^g - p_{\mathrm{DF}}^g)$, $(p_{\mathrm{GO}}^g - p_{\mathrm{GO}}^e)$, $(p_{\mathrm{GO}}^g - p_{\mathrm{DF}}^e)$, $(p_{\mathrm{DF}}^g - p_{\mathrm{GO}}^e)$, $(p_{\mathrm{DF}}^g - p_{\mathrm{DF}}^e)$ and $(p_{\mathrm{GO}}^e - p_{\mathrm{DF}}^e)$, as illustrated in the scatter plots reported in Figure A1 in Appendix A. In all cases it is possible to observe a clear and strong positive association between prices. In fact, we will later find out that there is evidence of integration in all six markets (that is,

6. In an empirical study, Salvo and Huse (2013) find substantial heterogeneity in how consumers might switch between gasoline and ethanol. According to the authors, about 20% of flexible-fuel motorists remain choosing gasoline even when this fuel is priced 20% (in energy-adjusted terms) above ethanol. A similar percentage is found when ethanol is priced 20% above gasoline. This evidence, however, only applies to a small portion of fuel consumers; thus, switching in response to price differentials appears to remain as a valid option for the majority who are not willing to pay as high a premium.

the price differentials are stationary). This extended approach is the one that we follow in our empirical analysis.

3. ECONOMETRIC MODELLING STRATEGY

The econometric modelling strategy that we follow in this paper starts employing tools from time-series analysis, but in a way that subsequently uses techniques from cross-section analysis. Commencing with time-series analysis, we apply the Pesaran (2007) pairwise approach to the study of gasoline and ethanol price differentials. More specifically, let us denote $p_{i,t}$ as the energy-equivalent price of fuel in market i at time t, where a "market" is now defined as a particular type of fuel (ethanol or gasoline) sold in a particular state (of which there are 27). Thus, there are a total of $N = 2 \times 27 = 54$ price series corresponding to each of these distinct fuel markets. Then, let $p_{ij,t} = p_{i,t} - p_{j,t}$ denote the price differential between any two markets $i \neq j$. It ought to be noticed that $p_{ij,t}$ permits the computation of all arbitrage opportunities that could exist, namely $(p_i^g - p_i^e)$, $(p_i^g - p_j^e)$, $(p_i^e - p_j^e)$ and $(p_i^g - p_j^g)$. The idea underlying the pairwise approach is to investigate the order of integration of all $(N(N-1)/2)$ price differentials with the purpose of identifying those that appear to behave as stationary processes (that is, those that are bounded).

For this, we employ the augmented Dickey and Fuller (1979) and the Leybourne (1995) unit root tests, which we refer to as ADF and $\text{ADF}_{\max}$, respectively. The ADF test regression that we specifically estimate, using ordinary least squares (OLS), is:

$$\Delta p_{ij,t} = \beta_1 + \beta_2 p_{ij,t-1} + \sum_{k=1}^{s_{\max}} \gamma_k \Delta p_{ij,t-k} + \varepsilon_{ij,t} \tag{1}$$

where Δ is the first difference operator, $s_{\max}$ is the number of lags of the dependent variable that are included in the test regression to account for residual serial correlation, and $t = 1, \ldots, T$ time observations. In this setting, the unit-root null is $H_0 : \beta_2 = 0$, against the alternative that the series is stationary around a constant mean, that is $H_0 : \beta_2 < 0$. Denoting the ADF test based on the regression t-statistic for $\beta_2 = 0$ in (1) as ADF_f, this value is compared against the critical values estimated by Cheung and Lai (1995) using response surfaces. As to the Leybourne test, consider now the reverse realisation of $p_{ij,t}$, given by $p^*_{ij,t} = p_{ij,T+1-t}$. The ADF-type regression applied to $p^*_{ij,t}$ is:

$$\Delta p^*_{ij,t} = \beta^*_1 + \beta^*_2 p^*_{ij,t-1} + \sum_{k=1}^{s_{\max}} \gamma^*_k \Delta p^*_{ij,t-k} + \varepsilon^*_{ij,t} \tag{2}$$

Let us denote ADF_r as the ADF test based on the regression t-statistic for $\beta^*_2 = 0$ in (2), which also includes $s_{\max}$ lags of the dependent variable to allow for residual serial correlation. Within this framework, Leybourne (1995) defines the $\text{ADF}_{\max}$ test as $\text{ADF}_{\max} = \max(\text{ADF}_f, \text{ADF}_r)$. The specific critical values that we use for inference are those estimated by Otero and Smith (2012) using response surfaces. The nominal size of the underlying unit root test statistic is α.

Defining z_{ij} as an indicator function that takes the value of one when the corresponding unit-root test statistic is rejected at significance level α, and zero otherwise, Pesaran (2007) considers the fraction of the $(N(N-1)/2)$ price differentials for which the unit-root null is rejected, and proposes a test statistic given by:

$$\bar{z}_{ij} = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} z_{ij} \tag{3}$$

Pesaran shows that although the underlying individual unit-root tests are not cross-sectional independent, under the null hypothesis of non-stationarity (that is, divergence), the expected value of $\bar{z}_{ij}$ is equal to the nominal size α as $N,T \rightarrow \infty$. In turn, support for the alternative of stationarity (that is, convergence) occurs whenever $\bar{z}_{ij} > \alpha$.[7] Pair-wise studies of price convergence, such as Yazgan and Yilmazkuday (2011) and Abbott and De Vita (2011), focus on the computation of the relative frequency of rejections, that is $\bar{z}_{ij}$. Even though the estimate of the fraction of rejections $\bar{z}_{ij}$ is in itself a parameter of interest, in this paper we focus instead on the examination of all individual stationary outcomes, that is $z_{ij} = 1$, in order to determine the variables, if any, that help to explain the speed at which prices adjust to reach their long-run equilibrium level.

In a sense, support for the finding that a price differential is stationary is equivalent to saying that the two prices are cointegrated with a known cointegrating vector, that is equal to $[1, -1]'$. The cointegrating vector plays the role of an attractor or long-run equilibrium relationship such that in the short run prices may deviate from it, but not by an ever growing amount since arbitrage opportunities are expected to act so as to restore equilibrium. Notice also that for non-stationary price differentials discrepancies with respect to the long-run equilibrium are not bounded. This implies that in these cases economic forces are not acting to restore equilibrium, possibly because of the presence of rigidities and other market frictions. In other words, starting from a point in which price pairs are cointegrated, if one state has prices different enough from another state, then someone will do the arbitrage. The interest of the paper is on how quickly that arbitrage happens.

At this stage of the analysis, it is worth noticing that the traditional approach to test for market integration involves computing price differentials with respect to a price that is chosen, sometimes arbitrarily, as a reference or benchmark price. However, this traditional approach utilises in a partial manner the information contained in the individual series, and therefore the results may be dependent on the chosen reference price. In sharp contrast to this, the pairwise approach exploits all the information contained in the series, and in doing so enhances the scope of the analysis. For instance, it may well occur that $(p_1^g - p_2^g)$ and $(p_1^g - p_3^e)$ are found to be non-stationary, so that the hypothesis of market integration does not hold for these price differentials, while $(p_2^g - p_3^e)$ is found to be stationary, and therefore market integration holds for this latter differential. This result could be explained by the existence of a highly persistent factor (or trend) that is common to both p_2^g and p_3^e, but is not shared by p_1^g (see Pesaran 2007; section 4).

Up to this point, our empirical analysis has relied on the examination of the time-series properties of the gasoline and ethanol prices under consideration. Next, we turn to the cross-section part of the analysis to understand the drivers that explain the speed of adjustment of the price differentials. To do this, we focus on the differentials that are stationary, and estimate the associated half-life of a shock, which we denote hl_{ij}. Notice that in the notation used for hl_{ij} we drop the subindex t since we focus on the price differentials that are stationary, and for these the mean and variance are constant through time. For the purposes of the cross-section analysis that follows, hl_{ij} is considered in logarithms and denoted lhl_{ij}.[8]

To specify the cross-section model for lhl_{ij} we follow Holmes, Otero, and Panagiotidis

7. For a recent test of cross-sectional dependence and the extent to which it occurs see Bailey, Kapetanios, and Pesaran (2015), who propose a method of measuring the extent of interconnections in large panels of data.

8. The half-life is approximated using the formula $-\ln(2)/(1 + \hat{\beta}_2)$, where $\hat{\beta}_2$ denotes the autoregressive coefficient in the corresponding ADF test regression (1); see Goldberg and Verboven (2005). We opt for using the half-life rather than $\hat{\beta}_2$ because the former can be intuitively interpreted as the time it takes for half the initial disequilibrium to be eliminated. The half-life is considered in logarithms in order to ease the interpretation of the regression coefficients of the cross-section analysis.

(2013), who consider a reduced form that includes exogenous cost or supply-side variables, demand-side variables and geographical variables as possible drivers. Cost or supply-side variables include the absolute value of the tax differential between states i and j, denoted $dtax_{ij}$.[9] The variable $dtax_{ij}$ aims to capture the effect of differentiated levels of taxation across products and/or states. The second cost or supply-side variable that we consider is the absolute differential in the logarithm of gas stations density, $dlgs_{ij} = |lgs_i - lgs_j|$, where lgs_i and lgs_j denote the logarithms of the number of gas stations (per 100,000 vehicles) in states i and j, respectively.[10] The variable $dlgs_{ij}$ seeks to measure fuel allocation effects between states. The third supply-side variable that we include is the absolute differential in sugarcane mills density, $dmill_{ij} = |mill_i - mill_j|$, where $mill_i$ and $mill_j$ denote the number of sugarcane mills (per million vehicles) in states i and j, respectively. The estimated coefficients on these three variables are expected to be positive, indicating lower speed of adjustment (that is, higher half-life) when the differentials in tax rates, gas stations density, and sugarcane mills density are higher.

To account for demand-side variables we include the absolute difference in the logarithm of the population density, denoted $dlpd_{ij} = |lpd_i - lpd_j|$, where lpd_i and lpd_j are the logarithms of the population densities in states i and j, respectively. The variable $dlpd_{ij}$ aims to capture city congestion effects within states as well as the geographic area of each state. Another variable is the absolute differential in the logarithm of real per-capita GDP in states i and j, which we denote $dlgdp_{ij}$ and it seeks to measure the population income. The coefficients associated to $dlpd_{ij}$ and $dlgdp_{ij}$ are expected to be positive supporting the view of lower speed of adjustment (that is, higher half-life) when the differentials in population densities and real per-capita output are higher.

Regarding geographical variables, we consider the logarithm of the distance between states i and j, which is denoted $ldist_{ij}$. The hypothesis of interest here is whether a longer distance is associated with a higher half-life and therefore a slower speed of adjustment back to equilibrium. Therefore, the coefficient associated to this variable is expected to be positive. Finally, we also include the dummy variable dug_{ij} which takes the value of one when the two prices in a differential refer to gasoline (and zero otherwise), and the dummy variable due_{ij} when they refer to ethanol (and zero otherwise). The role of these two dummy variables is to capture differentiated speeds of adjustment when prices correspond to the same fuel. The resulting cross-section regression model is:

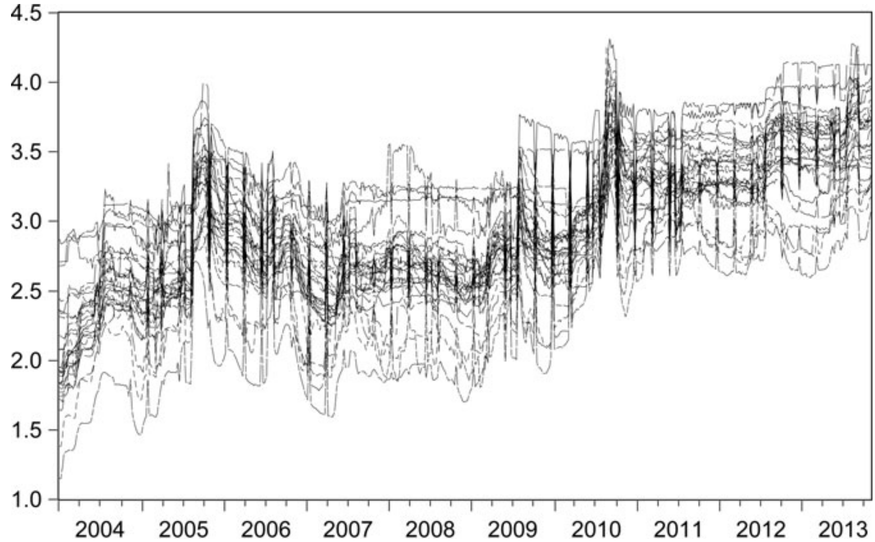
$$\begin{aligned} lhl_{ij} = \beta_1 + \beta_2 ldist_{ij} + \beta_3 dtax_{ij} + \beta_4 dlgs_{ij} + \beta_5 dmill_{ij} + \beta_6 dlpd_{ij} + \beta_7 dlgdp_{ij} \\ + \beta_8 dug_{ij} + \beta_9 due_{ij} + \epsilon_{ij} \end{aligned} \tag{4}$$

where, in addition to the variables already defined, ϵ_{ij} is the equation error term. The explanatory variables in the half-life regression enter as the absolute value of a differential because our analysis focuses mainly on the arbitrage opportunities offered by trading ethanol and gasoline within/between states, and the speed at which relative prices change. Therefore, the definition of the variables fits the purpose of testing whether the speed of adjustment can be explained by state similarities, which in turn are reflected in the magnitude of differentials in taxes, gas stations density, sugarcane mills density, population density, and real per-capita GDP, rather than the sign of these variables.

9. The variable $dtax_{ij}$ includes all possible tax differentials, that is $(tax_i^g - tax_i^e)$, $(tax_i^g - tax_j^e)$, $(tax_i^e - tax_j^e)$ and $(tax_i^g - tax_j^g)$, where tax_k^g and tax_k^e denote the tax rates applicable to gasoline and ethanol in state $k = i, j$, respectively.

10. In Brazil the overwhelming majority of gas stations are dual fuel, and so we assume that this variable is the same for both gasoline and ethanol.

Figure 1a: Post-tax Ethanol prices in levels in Brazilian states: 2004w19-2014w16



4. DATA AND EMPIRICAL ANALYSIS

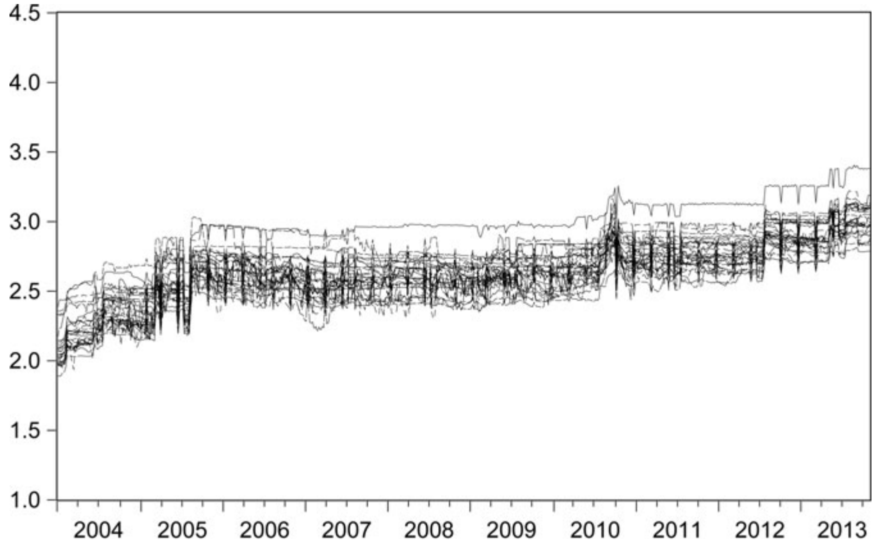
The database consists of the post-tax and pre-tax weekly average retail prices (measured in Brazilian reais, also denoted R$, per litre) of gasoline (gasolina comun) and hydrous ethanol (álcool comum) in each of the 27 Brazilian states[11] over the period 2004w19 to 2014w16, for a total of $T = 514$ time observations.[12] All the price series data are obtained from the Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP), which gathers the information from fuel stations in 411 Brazilian municipalities. For completeness, prices are analysed both in levels and also after applying the logarithmic transformation, so that in the latter case when price differentials are calculated the resulting values constitute approximate measures of the differentials in percentage terms. Overall, four cases will be considered, namely: i.) post-tax prices in logarithms; ii.) post-tax prices in levels; iii.) pre-tax prices in logarithms; and iv.) pre-tax prices in levels. Visual inspection of the plots of the price series over time (see Figures 1 and 2) along with some descriptive statistics (see Table 1) reveals that post-tax prices in levels for ethanol are more dispersed than post-tax gasoline prices, as the standard deviation of the former amounts to approximately R$ 0.33 while for the later it is R$ 0.22; this finding can be explained by the price control over gasoline.

In addition to the price series, to perform the subsequent cross-section analysis stated in equation (3), we consider the following information on the variables included as regressors in that equation (the corresponding summary statistics are reported in Table 1). First, to construct $dtax_{ij}$ one could ideally calculate the average (gasoline and ethanol) tax rates over the whole sample

11. The states of the Brazilian federation are: Acre, Alagoas, Amapá, Amazonas, Bahia, Ceará, Distrito Federal, Espírito Santo, Goiás, Maranháo, Mato Grosso, Mato Grosso do Sul, Minas Gerais, Pará, Paraíba, Paraná, Pernambuco, Piauí, Rio de Janeiro, Rio Grande do Norte, Rio Grande do Sul, Rondônia, Roraima, Santa Catarina, São Paulo, Sergipe and Tocantins.

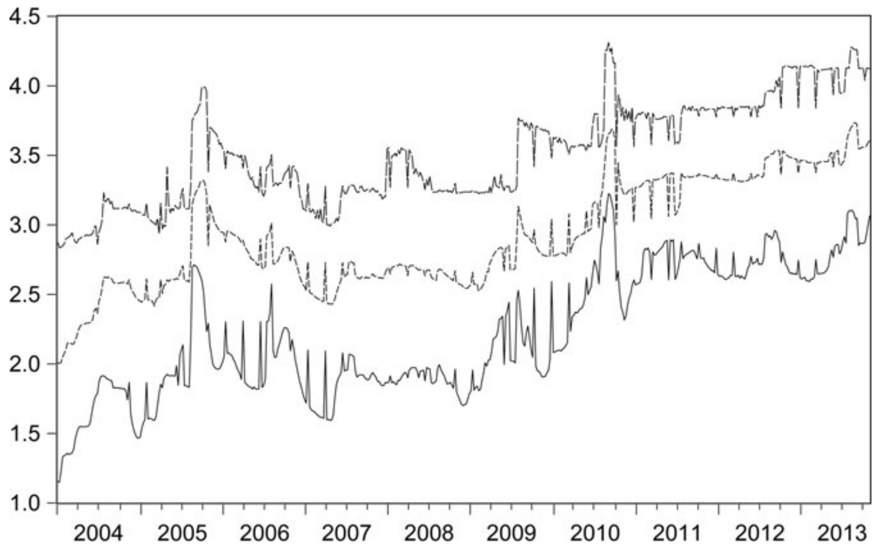
12. Unfortunately consumer fuel price series before taxes are not available from the ANP. However, to have an approximation to these price series for ethanol, we remove the corresponding tax rate in each state. As to gasoline, before subtracting the corresponding tax rate in each state, we calculate the price of pure gasoline and anhydrous ethanol by applying the blending mandate applicable in every period.

Figure 1b: Post-tax Gasoline prices in levels in Brazilian states: 2004w19-2014w16



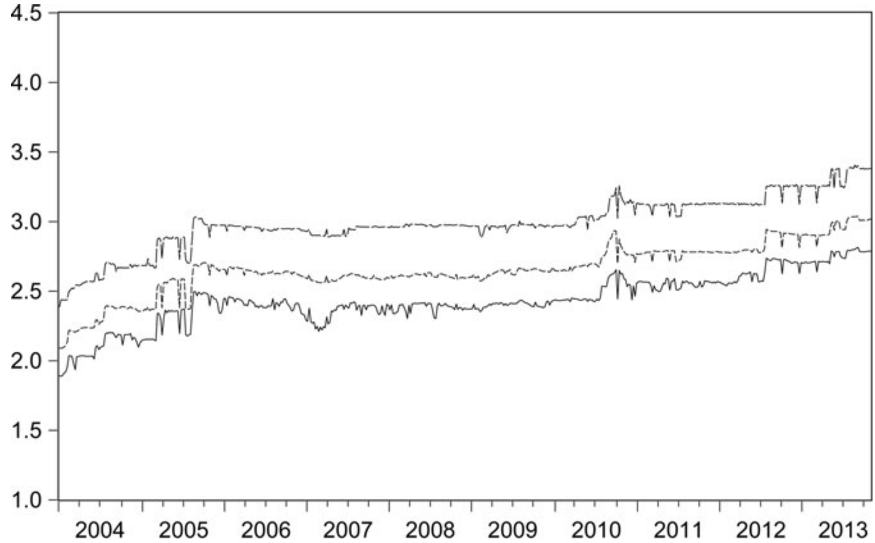
Note: Y-axis is in R$ per litre. Prices are in energy equivalents.

Figure 2a: Cross-sectional minimum, mean and maximum of post-tax Ethanol prices in Brazilian states: 2004w19-2014w16



period of interest, so that potential variations in tax levels over a longer time span are accounted for. However, data limitations prevent us from doing this, and so instead we employ average tax rates for the years after 2011; caution must be therefore exercised when interpreting the effect of the tax burden since taxation levels are not being evaluated over the whole sample period. The tax data was gathered from each state's Department of Finance. The mode of gasoline taxes is 25% and is applied by 16 out of the 27 states, including São Paulo, which is the state with the largest fuel consumption. In the case of ethanol taxes, the mode is 25% too, but São Paulo state applies a

Figure 2b: Cross-sectional minimum, mean and maximum of post-tax gasoline prices in Brazilian states: 2004w19-2014w16



Note: Y-axis is in R$ per litre. Prices are in energy equivalents.

Table 1: Descriptive Statistics

Variable	Mean	Standard Deviation	Min	Max
Pre-tax ethanol price (R$/litre)	2.00	0.26	0.76	3.17
Pre-tax gasoline price (R$/litre)	1.95	0.14	1.44	2.67
Post-tax ethanol price (R$/litre)	2.92	0.33	1.15	4.31
Post-tax gasoline price (R$/litre)	2.66	0.22	1.89	3.41
State ethanol tax	24.3%	3.2%	12.0%	29.0%
State gasoline tax	26.0%	1.4%	25.0%	30.0%
Fuel stations per 100,000 vehicles	77.91	23.36	26.53	124.03
Sugarcane mills	7.05	11.45	0.00	58.02
Population density (Inhabitants/km2)	64.14	98.84	1.74	407.68
GDP per capita (R$)	19,643.5	12,256.4	7,745.8	68,360.7

rate of 12%. Second, for $dlgs_{ij}$ we use the ratio between the average number of gas stations in each state over the period 2004-2013, which was calculated using data from the Anuário Estatístico Brasileiro do Petróleo, Gás Natural e Biocombustíveis of the ANP, and the corresponding total number of 100,000 vehicles during the same period, where information on the latter variable is taken from Anuário of the Departamento Nacional de Trânsito. Third, for $dmill_{ij}$ we averaged the number of sugarcane (ethanol and mixed) mills (per million vehicles) in each state during the years 2007-2014. Information was gathered from the report of registered sugarcane mills (Relaçao de Instituições Cadastradas No Departamento de Cana-de-Açúcar e Agroenergia) in the Ministerio da Agricultura Pecuaria e Abastecimento. Alagoas is by far the state with the highest number of sugarcane mills (per million vehicles) with 58 mills, followed by Mato Grosso do Sul with 20 mills, and Goiás and Paraíba each of which with 13 mills. Only four states (namely Amapá, Distrito Federal, Roraima and Santa Catarina) do not count with sugarcane mills. Fourth, for $dlpd_{ij}$ we use average population in the years 2000, 2007 and 2010 along with the corresponding area in each state, both of which are taken from the Instituto Brasileiro de Geografia e Estatística. For real per-capita GDP at the state level, nominal values over the period 2004-2011 are deflated using the consumer price index, averaged over time, and divided by the population figures described earlier. Finally, for the geographical distance between states i and j, $dist_{ij}$, we calculate the shortest great-circle distance in kilometres between centroids of any pair of states, where information on centroids is obtained through a Geographical Information System. To assess the robustness of the results to alternative measures of distance, we also consider the shortest road distance between the main cities of any pair of states. The results obtained using the two measures of distance are qualitatively the same.

Table 2 reports the percentage of rejections of both the ADF and $\mathrm{ADF}_{\max}$ unit root tests based on all the differentials that can be constructed when using gasoline prices (i.e. 351), ethanol prices (i.e. 351), cross ethanol-gasoline prices between states (i.e. 702), cross ethanol-gasoline prices within states (i.e. 27), and all gasoline and ethanol prices (i.e. 1,431) for the four cases under consideration. The unit root test regressions include an intercept, and the order of augmentation is determined using the Akaike information criterion (AIC), with $s_{\max} = 6$ lags (results were very similar when $s_{\max}$ was set at four and eight lags). Inference is performed at the $\alpha = 0.05$, 0.10 significance levels. In general terms, in all cases the percentage of rejections exceeds the underlying size of the unit root test statistics. Looking first at the differentials that involve all post-tax gasoline and ethanol prices in logarithms, the rejection frequency reaches 56.9% (69.5%) at the $\alpha = 0.05$ (0.10) significance level. When examining post-tax prices in levels (case ii) the results are quite similar to those when prices are in logs. Regarding pre-tax prices in logs (case iii), most of the proportions of rejection of the unit root tests increase; for example, under the ADF at the $\alpha = 0.10$ level the percentage of rejections for all pairs is 80%, which is significantly higher than in the previous two cases. However, when pre-tax prices are measured in levels, the proportions of rejections fall with respect to the other three cases; for instance, under the ADF at $\alpha = 0.05$, the percentage of rejections is 42.4%, the smallest percentage among all cases. When unit-root inference is based on the $\mathrm{ADF}_{\max}$ test the percentages of rejection for post-tax prices in logarithms are smaller (except when only gasoline price differentials are involved), while for prices in levels most of the percentages of rejection are higher. Finally, when both fuels are considered separately, ethanol prices are more integrated than gasoline prices mainly due to the price control policy over the latter fuel (for instance, the ADF test yields a rejection frequency of 59.8% (70.9%) at the $\alpha = 0.05$ (0.10) significance level for post-tax gasoline prices in logarithms, while 86.0% (92.5%) also at the $\alpha = 0.05$ (0.10) significance level for post-tax ethanol prices in logarithms). To summarise, the previous findings not only reveal the large extent of integration within the ethanol and gasoline markets, but also between them. Moreover, the results highlight the importance of allowing for spatial effects when analysing fuel prices in Brazil.[13]

13. In an additional analysis, the prices for ethanol and gasoline were also studied after removing the purely time-series component, by regressing each price series on a set of week dummy variables, and analysing the resulting residuals for each state separately. The idea behind is to determine whether the stationarity in price differentials is coming from the time-series versus cross-sectional component (or both), and what this implies about the integration of fuel markets in Brazil. Results not reported here indicate that for 51 (out of 54) price series the joint null hypothesis that the estimated coefficients on the week dummy variables are equal to zero cannot be rejected. As a result, the percentages of rejection of the unit-root null for the price differentials are qualitatively similar to those presented in Table 2. For example, using the ADF test at the 10% level the proportion of stationary ethanol-gasoline price differentials after removing the effect of the week dummy variables is 55.7% (48.1%) between (within) states. The respective percentages in Table 2 are 57.8% and 48.1%. In conclusion, the imposition of the same pattern of (deterministic) seasonality across all states does not appear to affect the integration properties of fuel markets.

Table 2: Proportion of Stationary Gasoline and Ethanol Price Differentials

Unit root test	α	Only Ethanol	Only Gasoline	Gasoline and Ethanol: Between states	Gasoline and Ethanol: Within states	All
Pairs		351	351	702	27	1,431
			Post-tax prices in logarithms			
ADF	5%	86.0%	59.8%	41.5%	37.0%	56.8%
ADF	10%	92.5%	70.9%	57.9%	48.1%	69.4%
ADF_{max}	5%	70.3%	62.1%	36.7%	37.0%	51.2%
ADF_{max}	10%	73.7%	74.3%	48.8%	40.7%	61.0%
			Post-tax prices in levels			
ADF	5%	83.1%	57.5%	34.0%	25.9%	51.7%
ADF	10%	89.4%	68.0%	43.4%	44.4%	60.7%
ADF_{max}	5%	76.6%	64.3%	36.6%	40.7%	53.3%
ADF_{max}	10%	84.0%	81.0%	50.7%	48.1%	66.3%
			Pre-tax prices in logarithms			
ADF	5%	91.7%	34.7%	62.5%	59.2%	62.8%
ADF	10%	97.1%	53.8%	84.0%	96.2%	80.0%
ADF_{max}	5%	82.0%	49.8%	17.2%	14.8%	41.0%
ADF_{max}	10%	87.7%	67.8%	26.9%	22.2%	51.5%
			Pre-tax prices in levels			
ADF	5%	87.4%	29.9%	26.6%	29.6%	42.4%
ADF	10%	92.3%	45.2%	46.8%	48.1%	57.6%
ADF_{max}	5%	89.1%	49.9%	24.9%	18.5%	46.4%
ADF_{max}	10%	93.4%	65.2%	43.3%	40.7%	60.9%

Notes: The underlying unit-root test regressions include a constant, and the number of lags of the dependent variable that are included in the test regression is selected using the Akaike information criterion with $s_{max} = 6$ lags. Both the ADF and ADF_{max} tests are performed at significance level α. The critical values of ADF and ADF_{max} tests are calculated using the response surfaces estimated by Cheung and Lai (1995) and Otero and Smith (2012), respectively.

Figures A2 to A5 in Appendix A display the relative number of stationary price differentials in each state, based on the ADF unit root test at the 5% significance level, for the four cases already presented. In these figures, the grey scale reveals the high degree of integration, where any state has at least 25% stationary price differentials (out of a total of 106 possible pairs). Also, these figures indicate that the highest degree of fuel market integration occurs in the states located in the central part of the country (e.g. Goiás and DF).

Computation of the half-life for the differentials that turn out to be stationary (based on the ADF test at the 5% significance level) shows that under all cases the average half-life for gasoline is slightly smaller than that for ethanol, that is 6.4 weeks compared to 7.3 weeks. When considering post-tax prices in levels, the estimated speeds of adjustment for the two fuels tend to be somewhat similar, with average half-lives of 6.3 weeks for gasoline and 6.7 weeks for ethanol. Interestingly, when using pre-tax prices the corresponding half-lives are significantly higher than those obtained when using post-tax prices (for the prices in logarithms the average half-lives for gasoline and ethanol pre-tax prices are respectively 11.3 and 12.8 weeks, while for the prices in levels the corresponding values are 10.6 for gasoline and 12.1 for ethanol). Overall, the quicker speed of adjustment in gasoline prices can be explained by the larger pipeline network and infrastructure to transport it, whereas ethanol is mostly dependent on road transportation. Furthermore, these results also show that regional tax differentials appear to play an important role in the speed of adjustment of fuel prices.

Next, we attempt to understand the factors that explain the speed of adjustment of gasoline and ethanol price differentials in Brazil. To do this, we focus on the cases where the null of non-stationarity is rejected using the ADF test at the 5% significance level (qualitatively similar results are obtained when using a 10% significance level). Ordinary least squares estimation of equation (3) for the four cases yields the results reported in the column labelled model I in Tables 3 to 6. Starting with the results for post-tax prices in logarithms in Table 3, model I, the estimated coefficients on $ldist_{ij}$, $dtax_{ij}$, $dmill_{ij}$, $dlgs_{ij}$, $dlgdp_{ij}$ and $dlpd_{ij}$ are positive, as expected, although those for the last three variables are not statistically different from zero.[14] These significant coefficients indicate that the half-life of shocks is higher (i.e. the speed of adjustment is slower) for states that are farther apart, for states that are more dissimilar in terms of fuel taxation, and for states with larger differential in sugarcane mills density, respectively. For case (ii) (post-tax prices in levels), the intercept and the coefficients on the dummy variables due_{ij} and dug_{ij} result highly statistically significant, while the estimated coefficient on distance has the expected positive sign as is marginally statistically significant when using a one-tailed test at the 10% significance level; the remaining coefficients are not statistically different from zero (see Table 4). Turning to Table 5, when applying the logarithmic transformation to pre-tax prices the coefficients for $dtax_{ij}$, $dlgdp_{ij}$ and $dlgs_{ij}$ are significant; here, the first two are positive and the third one is negative. Finally, when using pre-tax prices in levels only the coefficients for $ldist_{ij}$, $dtax_{ij}$ and $dlgs_{ij}$ are significant (see Table 6). In summary, the coefficients on state tax differentials across cases turn out to be positive and statistically significant; furthermore, those under pre-tax prices are substantially higher than those under post-tax prices, which reinforce the previous hypothesis that regional tax differentials increase the speed of adjustment (the smaller the coefficient the smaller the effect on increasing the half-life). Similarly, distance appears to play a role in explaining the speed of adjustment of fuel prices.

Excluding the insignificant regressors in model I yields model II in Tables 3 to 6, where the remaining estimated coefficients are virtually unchanged compared to those reported in model

14. It is interesting to examine the results when the half-lives of all 1,431 price differentials are used to estimate equation (3), so that one is implicitly assuming that the inability to reject the unit-root null might not necessarily imply an "invalid estimate" of the half-life. Results not reported here indicate that the estimated model is inferior to those reported in Table 3 as model I and II, in the sense that the only coefficients that are statistically different from zero are those for the ethanol and gasoline dummy variables. Alternatively, one could have also adopted the view that the issue is not that the observations with $\hat{\beta}_2$ close to zero are not informative, but that the half-life statistic has $\hat{\beta}_2$ in the denominator. Thus, in additional estimations we also estimated the half-life regression for all pairs with $\hat{\beta}_2 < -0.05$, as well as when $\hat{\beta}_2$ is used as dependent variable instead of the half-life. However, compared to the results in Table 3, the results of these alternative estimations are once again inferior. In conclusion, the weakening of the findings when half-lives associated to non-stationary price differentials are employed for estimating equation (3) appears consistent with the fact that for non-stationary price differentials there are no forces that act in order to restore equilibrium.

Table 3: Determinants of the Half-life of Fuel Post-tax Price Differentials in Logarithms

Model	I: All pairs		II: All pairs		III: Only ethanol		IV: Only gasoline		V: Ethanol-gasoline between states	
Variable	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)
Intercept	2.18***	(0.041)	2.17***	(0.024)	1.91***	(0.031)	1.73***	(0.051)	2.12***	(0.028)
ldist	0.06***	(0.022)	0.065**	(0.021)	0.044	(0.031)	0.165**	(0.046)	-0.019	(0.034)
dtax	0.776**	(0.359)	0.764**	(0.362)	-0.521	(0.566)	-3.46***	(2.624)	2.41***	(0.366)
dlgdp	0.007	(0.035)								
dlpd	0.008	(0.011)								
dlgs	-0.009	(0.014)								
dmill	0.001*	(0.001)	0.001*	(0.001)	0.000	(0.001)	0.004*	(0.002)	0.005*	(0.001)
due	-0.32***	(0.029)	-0.32***	(0.029)						
dug	-0.45***	(0.036)	-0.45***	(0.036)						
Obs.	814		814		302		210		302	
R^2	0.214		0.213		0.009		0.074		0.095	
Normal	0.869	[0.647]	0.718	[0.698]	9.551	[0.008]	3.508	[0.173]	3.554	[0.169]
Hetero	1.485	[0.027]	1.754	[0.029]	0.526	[0.855]	0.443	[0.910]	1.998	[0.039]
RESET	0.012	[0.911]	0.063	[0.801]	0.014	[0.707]	0.021	[0.882]	3.032	[0.082]
Moran I	18.080	[0.000]	18.090	[0.000]	11.316	[0.000]	9.608	[0.000]	16.770	[0.000]

Notes: The dependent variable is in logarithms. Standard errors (in parentheses) are heteroskedasticity consistent. Model I includes all observations for which price differentials are stationary based on the ADF test at the 5% level. Model II excludes insignificant variables in Model I. Model III excludes insignificant variables in Model I and uses for estimation only-ethanol observations. Model IV excludes insignificant variables in Model I and uses for estimation only-gasoline observations. Model V excludes insignificant variables in Model I and uses for estimation only ethanol-gasoline observations between states. Normal is the Jarque and Bera (1987) χ_2^2 test for normality, which measures the difference of the skewness and kurtosis of the residuals with those from the standard normal distribution. Hetero is the *F*-version of the White (1980) test for heteroskedasticity, based on an auxiliary regression of the squared residuals on the original regressors, all their squares and cross products. RESET is the *F*-version of the Ramsey (1969) regression specification error test, based on an auxiliary regression of the dependent variable on the original regressors and the second power of the fitted values from the original regression. Moran I is the standardised version of the Moran (1950) test for spatial autocorrelation (which follows a standard normal distribution), using as weights a row standardized queen contiguity matrix. Numbers in squared brackets indicate the probability values of the test statistics.

* Significant at the 10% level; ** significant at the 5% level; ***significant at the 1% level.

Table 4: Determinants of the Half-life of Fuel Post-tax Price Differentials in Levels

Model	I: All pairs		II: All pairs	
Variable	Coeff.	(s.e.)	Coeff.	(s.e.)
Intercept	2.23***	(0.042)	2.23***	(0.019)
ldist	0.027	(0.022)		
dtax	-0.065	(0.378)		
dlgdp	-0.018	(0.034)		
dlpd	0.001	(0.011)		
dlgs	-0.001	(0.013)		
dmill	0.001	(0.001)		
due	-0.39***	(0.029)	-0.39***	(0.028)
dug	-0.48***	(0.035)	-0.48***	(0.035)
Obs.	740		740	
R^2	0.253		0.250	
Normal	8.077	[0.017]	8.363	[0.015]
Hetero	1.644	[0.007]	10.54	[0.000]
RESET	8.400	[0.003]	N/A	[N/A]
Moran I	17.506	[0.000]	17.566	[0.000]

Notes: The dependent variable is in logarithms. Standard errors (in parentheses) are heteroskedasticity consistent. Model I includes all observations for which price differentials are stationary based on the ADF test at the 5% level. Model II excludes insignificant variables in Model I. Models III-V in Table 3 cannot be estimated in this case. Normal is the Jarque and Bera (1987) χ^2_2 test for normality, which measures the difference of the skewness and kurtosis of the residuals with those from the standard normal distribution. Hetero is the *F*-version of the White (1980) test for heteroskedasticity, based on an auxiliary regression of the squared residuals on the original regressors, all their squares and cross products. RESET is the *F*-version of the Ramsey (1969) regression specification error test, based on an auxiliary regression of the dependent variable on the original regressors and the second power of the fitted values from the original regression. Moran I is the standardised version of the Moran (1950) test for spatial autocorrelation (which follows a standard normal distribution), using as weights a row standardized queen contiguity matrix. Numbers in squared brackets indicate the probability values of the test statistics.
* Significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

I.[15] Overall the diagnostic test statistics suggest that the resulting model is adequate (see the notes in Table 3-6 for details on the construction of the diagnostic tests). The Jarque-Bera test (Jarque and Bera 1987) for normality indicates that most residuals of the models do not depart from the standard normal distribution. The Ramsey regression specification error test (Ramsey 1969) does not detect general functional form misspecification for most of the cases. However, the White test

15. Although it is possible to argue that the differential in gas stations density is potentially endogenous, the estimated coefficient on this variable is not statistically different from zero, and therefore is not likely to challenge our findings regarding distance and tax rate differentials. Moreover, it is difficult to establish empirically whether the speed of adjustment of gasoline–ethanol, gasoline–gasoline and ethanol–ethanol relative prices affect the differential in the number of fuel stations, because the majority of them are dual fuel. In this sense, perhaps it is more likely to expect that relative prices affect the number of pumps for each fuel within each station. Unfortunately, we do not have this information for the purposes of our analysis.

Table 5: Determinants of the half-life of fuel pre-tax price differentials in logarithms

Model	I: All pairs		II: All pairs		III: Only ethanol		IV: Only gasoline		V: Ethanol-gasoline between states	
Variable	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)
Intercept	3.41***	(0.036)	3.41***	(0.033)	2.59***	(0.069)	2.33***	(0.108)	3.37***	(0.039)
ldist	0.005	(0.021)								
dtax	3.36***	(0.339)	3.36***	(0.337)	3.29***	(0.533)	-1.269	(2.202)	3.83***	(0.370)
dlgdp	0.066**	(0.0325)	0.063**	(0.032)	0.128*	(0.066)	0.020	(0.079)	0.037	(0.037)
dlpd	-0.011	(0.011)								
dlgs	-0.03***	(0.012)	-0.03***	(0.011)	-0.09***	(0.024)	0.012	(0.038)	-0.022	(0.013)
dmill	0.001	(0.000)								
due	-0.96***	(0.028)	-0.95***	(0.028)						
dug	-1.01***	(0.035)	-1.01***	(0.035)						
Obs.	899		899		322		122		455	
R^2	0.646		0.646		0.103		0.005		0.098	
Normal	11.13	[0.003]	10.56	[0.005]	1.745	[0.417]	0.136	[0.933]	12.69	[0.001]
Hetero	1.976	[0.000]	1.754	[0.029]	0.806	[0.610]	0.881	[0.544]	3.112	[0.001]
RESET	1.613	[0.204]	2.811	[0.000]	0.715	[0.398]	4.105	[0.045]	0.889	[0.346]
Moran I	22.165	[0.000]	22.230	[0.000]	12.116	[0.000]	6.827	[0.000]	24.480	[0.000]

Notes: The dependent variable is in logarithms. Standard errors (in parentheses) are heteroskedasticity consistent. Model I includes all observations for which price differentials are stationary based on the ADF test at the 5% level. Model II excludes insignificant variables in Model I. Model III excludes insignificant variables in Model I and uses for estimation only-ethanol observations. Model IV excludes insignificant variables in Model I and uses for estimation only-gasoline observations. Model V excludes insignificant variables in Model I and uses for estimation only ethanol-gasoline observations between states. Normal is the Jarque and Bera (1987) χ^2_2 test for normality, which measures the difference of the skewness and kurtosis of the residuals with those from the standard normal distribution. Hetero is the *F*-version of the White (1980) test for heteroskedasticity, based on an auxiliary regression of the squared residuals on the original regressors, all their squares and cross products. RESET is the *F*-version of the Ramsey (1969) regression specification error test, based on an auxiliary regression of the dependent variable on the original regressors and the second power of the fitted values from the original regression. Moran I is the standardised version of the Moran (1950) test for spatial autocorrelation (which follows a standard normal distribution), using as weights a row standardized queen contiguity matrix. Numbers in squared brackets indicate the probability values of the test statistics.

* Significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

Table 6: Determinants of the Half-life of Fuel Pre-tax Price Differentials in Levels

Model	I: All pairs		II: All pairs		III: Only ethanol		IV: Only gasoline		V: Ethanol-gasoline between states	
Variable	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)
Intercept	3.38***	(0.045)	3.33***	(0.041)	2.53***	(0.080)	2.17***	(0.115)	3.31***	(0.051)
ldist	0.050**	(0.026)	0.051**	(0.023)	0.12***	(0.032)	-0.024	(0.039)	-0.044	(0.041)
dtax	2.38***	(0.484)	2.38***	(0.489)	2.14***	(0.593)	3.702**	(1.885)	3.56***	(0.686)
dlgdp	0.065	(0.044)								
dlpd	-0.014	(0.013)								
dlgs	-0.03***	(0.014)	-0.03***	(0.012)	-0.07***	(0.024)	0.028	(0.033)	-0.014	(0.016)
dmill	0.000	(0.001)								
due	-0.94***	(0.029)	-0.94***	(0.032)						
dug	-1.01***	(0.040)	-1.00***	(0.039)						
Obs.	607		607		307		105		195	
R^2	0.606		0.604		0.083		0.030		0.100	
Normal	0.493	[0.781]	0.209	[0.900]	0.110	[0.946]	0.012	[0.993]	1.145	[0.563]
Hetero	2.095	[0.001]	2.503	[0.000]	1.214	[0.285]	1.317	[0.238]	2.162	[0.026]
RESET	0.185	[0.667]	0.085	[0.770]	6.369	[0.012]	0.707	[0.402]	2.536	[0.112]
Moran I	17.218	[0.000]	17.340	[0.000]	13.088	[0.000]	5.229	[0.000]	13.854	[0.000]

Notes: The dependent variable is in logarithms. Standard errors (in parentheses) are heteroskedasticity consistent. Model I includes all observations for which price differentials are stationary based on the ADF test at the 5% level. Model II excludes insignificant variables in Model I. Model III excludes insignificant variables in Model I and uses for estimation only-ethanol observations. Model IV excludes insignificant variables in Model I and uses for estimation only-gasoline observations. Model V excludes insignificant variables in Model I and uses for estimation only ethanol-gasoline observations between states. Normal is the Jarque and Bera (1987) χ^2_2 test for normality, which measures the difference of the skewness and kurtosis of the residuals with those from the standard normal distribution. Hetero is the *F*-version of the White (1980) test for heteroskedasticity, based on an auxiliary regression of the squared residuals on the original regressors, all their squares and cross products. RESET is the *F*-version of the Ramsey (1969) regression specification error test, based on an auxiliary regression of the dependent variable on the original regressors and the second power of the fitted values from the original regression. Moran I is the standardised version of the Moran (1950) test for spatial autocorrelation (which follows a standard normal distribution), using as weights a row standardized queen contiguity matrix. Numbers in squared brackets indicate the probability values of the test statistics.

* Significant at the 10% level; ** significant at the 5% level; ***significant at the 1% level.

(White 1980) for heteroskedasticity is rejected at the 5% significance level for some models, and so it seems prudent to perform inference based on White's heteroskedasticity consistent standard errors, which are reported in parentheses next to each regression coefficient. Finally, Moran's I test (Moran 1950) for spatial dependence in the residuals reject the null hypothesis that there is no spatial correlation for most of the models. Hence, spatial error regressions are performed for model I in all the four cases, and their results confirm the robustness of our estimates in terms of the magnitude and significance of the coefficients, which remain virtually unchanged as exhibited in Table 7.[16]

In models I and II the presence of the dummy variables due_{ij} and dug_{ij} implies that the intercept term must be interpreted as the half-life for the group against which comparisons are made, namely price differentials that involve prices of different fuels. Given that in all cases the estimated coefficients on due_{ij} and dug_{ij} are negative and statistically different from zero, with the former being smaller (in absolute value) than the latter, adjustment is quicker for price differentials that involve ethanol prices than it is for prices of different fuels, and even quicker for those that involve gasoline prices.[17]

Tables 3 to 6 also presents the results when model II is estimated using only stationary differentials for ethanol prices (model III), for gasoline prices (model IV), and for ethanol-gasoline prices between states (model V).[18] Here the idea is to allow for separate coefficients on each of the variables, rather than simply level shifts (which are accounted for with the ethanol and gasoline dummy variables). As can be seen, some of the coefficients become insignificant, and the magnitude of those that remain statistically significant remain largely unaffected. The loss of significance suggests that a model that imposes homogenous effects may not be valid.

5. CONCLUDING REMARKS

This paper studies integration in the gasoline and ethanol markets in Brazil, a large country with an established biofuel market, where currently most of the light-duty vehicles can use indifferently either ethanol or gasoline at any proportion. The distinguishing feature of the paper is that we explicitly incorporate spatial considerations into the analysis. We consider that these are very relevant in a country such as Brazil, where geographical conditions are complex and where tax regimes vary considerably across states. To test for spatial integration we adopt a time-series pair-

16. More specifically, the spatial error regression version of equation (3) that we estimate is $lhl = X\beta + \varepsilon$, where $\varepsilon = \lambda W + \upsilon$ and the error term υ is *iid*. Here, *lhl*, ε and υ are vectors of dimension 1,431 × 1, *X* is the matrix of regressors of dimension 1,431 × *k*, with *k* being the number of regressors, β is a $k \times 1$ vector of coefficients, *W* is the matrix of spatial weights of dimension 1,431 × 1,431, and λ is the spatial correlation parameter. It is worth noticing that because the model in equation (3) represents the spatial relations between pairs of states, it is not a traditional spatial model, but a model more related to a gravity equation of the type encountered in international trade. Therefore, the construction of the matrix of weights, which represents the relationships among 1,431 pairs, deserves special attention. In particular, following LeSage and Pace (2008), we construct *W* by first building a row standardized queen matrix of dimensions 27 × 27. This matrix gives a value of 1 to each of the contiguous state neighbors, including a value of 1 for the same state, and then each cell is weighted by the total sum of the row. The matrix is then quadruplicated to include the two fuels, resulting in a 54 × 54 matrix, but then the cases when the fuel is the same within a state are disregarded. The resulting matrix is replicated 27 times diagonally down-right to include all the relationships among states, resulting in a block-diagonal matrix of 1,431 × 1,431.

17. When the cross-section model II is estimated using price series that have been previously adjusted using a set of week dummy variables, results not reported here indicate that the estimated coefficients are qualitatively similar to those reported in Table 3. For example, the estimates of the coefficients on distance and taxes are 0.070 and 0.693, respectively, compared to 0.060 and 0.776 in Table 3.

18. The model that uses only the stationary price differentials for ethanol-gasoline within states is not estimated.

Table 7: Determinants of the Half-life for All Price Differentials (model I): Spatial Error Regression

Model	Fuel post-tax price differentials in logarithms		Fuel post-tax price differentials in levels		Fuel pre-tax price differentials in logarithms		Fuel pre-tax price differentials in levels	
Variable	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)
Intercept	2.18***	(0.046)	2.23***	(0.047)	3.40***	(0.041)	3.36***	(0.050)
ldist	0.063**	(0.026)	0.026	(0.025)	0.025	(0.025)	0.064**	(0.029)
dtax	0.741*	(0.444)	-0.259	(0.425)	2.89***	(0.416)	2.19***	(0.481)
dlgdp	0.032	(0.039)	0.000	(0.038)	0.066**	(0.036)	0.029	(0.045)
dlpd	0.003	(0.013)	-0.002	(0.012)	-0.014	(0.011)	-0.013	(0.014)
dlgs	-0.011	(0.014)	-0.003	(0.014)	-0.03**	(0.012)	-0.031**	(0.015)
dmill	0.002**	(0.001)	0.001	(0.001)	0.000	(0.001)	-0.001	(0.001)
due	-0.32***	(0.042)	-0.39***	(0.029)	-0.95***	(0.023)	-0.92***	(0.031)
dug	-0.45***	(0.046)	-0.48***	(0.040)	-1.01***	(0.053)	-0.99***	(0.055)
λ	0.33***	(0.003)	0.33***	(0.038)	0.46***	(0.030)	0.39***	(0.038)

Notes: The dependent variable is in logarithms. Standard errors in parentheses. The spatial error regression models were estimated using maximum likelihood. The spatial weight matrix is based on a row standardized queen contiguity matrix, which is constructed as described in footnote 16. The spatial correlation parameter is λ.

* Significant at the 10% level; ** significant at the 5% level; ***significant at the 1% level.

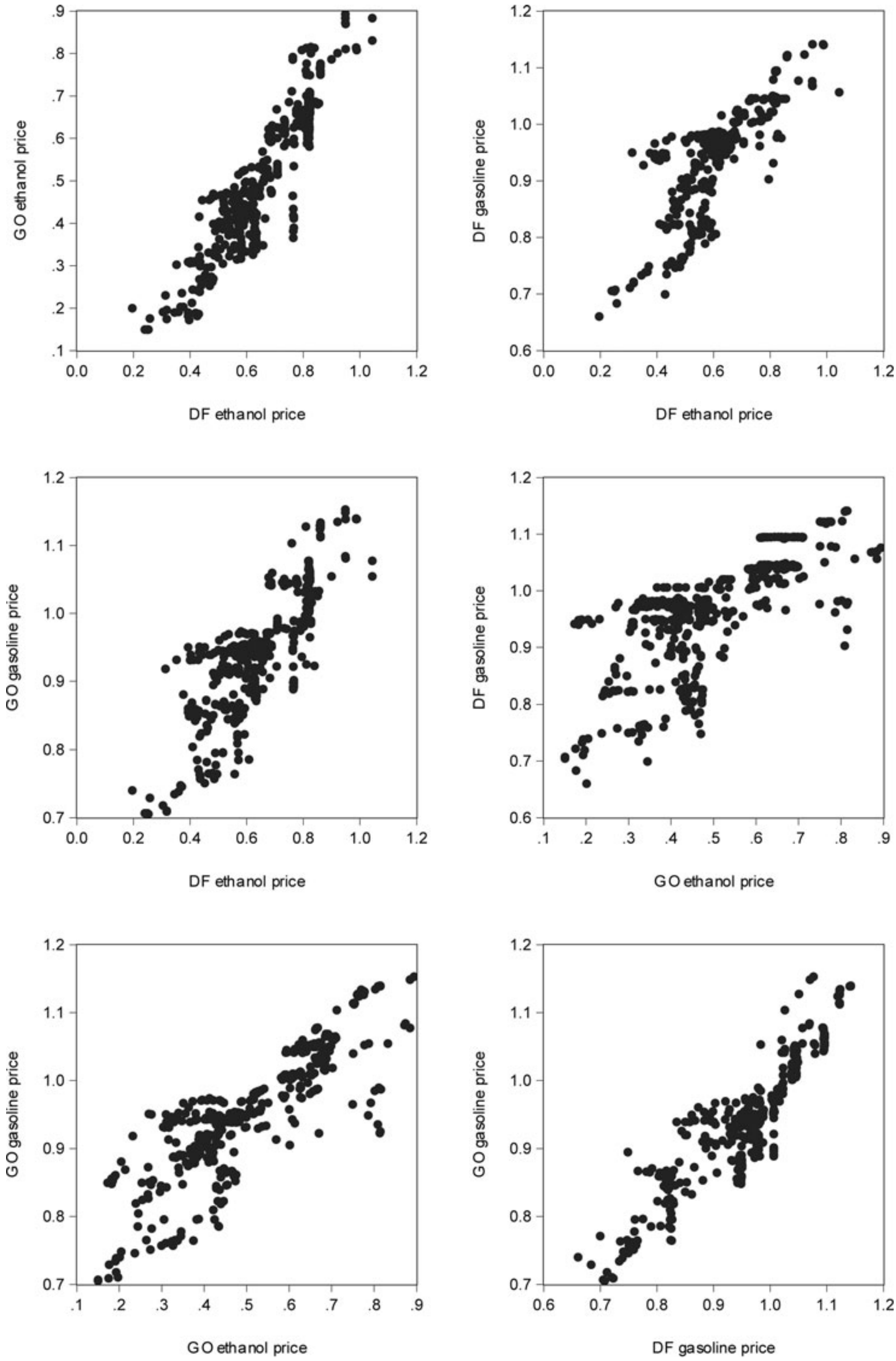
wise approach, but we also employ information from a cross-section approach. The pairwise view allows us to determine the proportion of stationary price differentials out of all the possible price differentials that can be constructed both within and between ethanol and gasoline for all 27 Brazilian states.

We present several important findings. Firstly, more than half of the fuel price differentials are stationary, which reveals the importance of accounting for spatial effects in the analysis. Second, the average half-life for gasoline price differentials is shorter than that for hydrous ethanol, which can be due to the larger pipeline network and infrastructure to transport gasoline, while ethanol is mostly dependent on road transportation. Thirdly, distance, the differentials in mills density and taxes play a role in determining the speed of adjustment of price differentials. Regarding distance, there is evidence that the longer the distance between two states, the slower the speed of adjustment back to equilibrium. With respect to the differential in sugarcane mills density, it has a positive effect implying a decrease in the speed of price adjustment, although the effect appears to be small in magnitude. As to taxes, the larger the existing differential in tax rates, that is between states and/or between ethanol and gasoline, the slower the speed of adjustment. Lastly, in comparison to price differentials that involve gasoline and ethanol prices, adjustment is quicker for price differentials that only involve ethanol prices, and even quicker for those that involve gasoline prices alone.

From the point of view of economic policy, our findings highlight the role played by factors that may obstruct fuel market integration between and/or within states, and also impede prices to adjust quickly and homogenously within the country. Bearing in mind the known defective conditions of many roads in the country, investment in better fuel transportation infrastructure, such as pipeline network projects to take ethanol/gasoline to the main consumer centres, can help to increase the extent of fuel market integration and reduce the time it takes for the corresponding prices to adjust. Finally, our results also illustrate how the speed of adjustment of prices to their long-run equilibrium value may be affected by altering the fuel taxation levels of a particular state; this is important because it is very possible for the optimal second-best tax to differ across states.

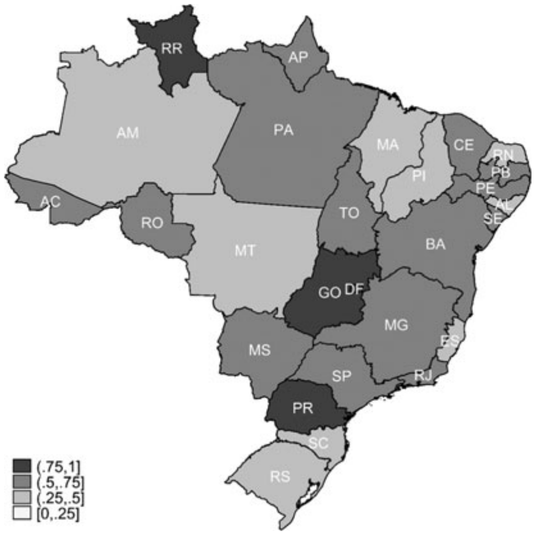
APPENDIX A

Figure A1: Scatter Plots of Gasoline and Ethanol Prices in DF and Goiás (GO)



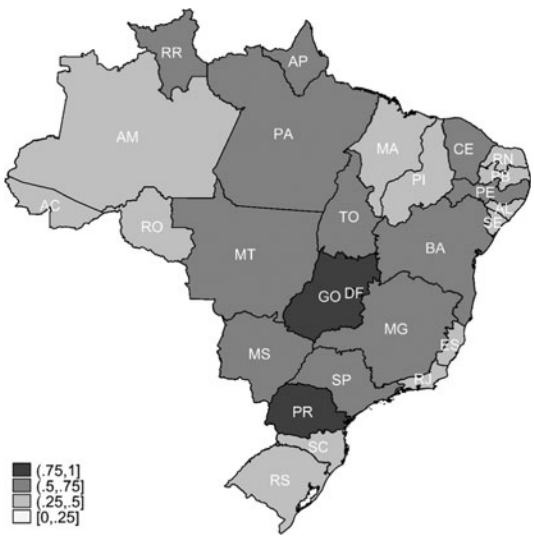
Note: Prices are in logarithms.

Figure A2: Stationary Fuel Post-tax Price Differentials in Logarithms



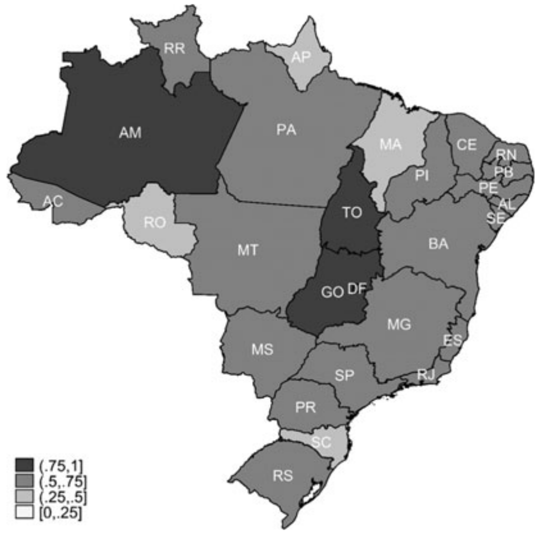
Note: The grey scale indicates the proportions of stationary price differential relationships for each state, based on the ADF unit root test at the 5% significance level. State abbreviations are: Acre (AC), Alagoas (AL), Amapá (AP), Amazonas (AM), Bahia (BA), Ceará (CE), Distrito Federal (DF), Espírito Santo (ES), Goiás (GO), Maranhão (MA), Mato Grosso (MT), Mato Grosso do Sul (MS), Minas Gerais (MG), Pará (PA), Paraíba (PB), Paraná (PR), Pernambuco (PE), Piauí (PI), Rio de Janeiro (RJ), Rio Grande do Norte (RN), Rio Grande do Sul (RS), Rondônia (RO), Roraima (RR), Santa Catarina (SC), São Paulo (SP), Sergipe (SE) and Tocantins (TO).

Figure A3: Stationary Fuel Post-tax Price Differentials in Levels



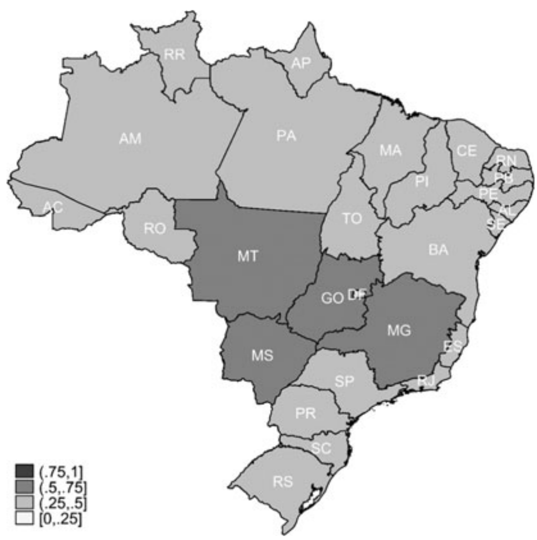
Note: The grey scale indicates the proportions of stationary price differential relationships for each state, based on the ADF unit root test at the 5% significance level. State abbreviations are: Acre (AC), Alagoas (AL), Amapá (AP), Amazonas (AM), Bahia (BA), Ceará (CE), Distrito Federal (DF), Espírito Santo (ES), Goiás (GO), Maranhão (MA), Mato Grosso (MT), Mato Grosso do Sul (MS), Minas Gerais (MG), Pará (PA), Paraíba (PB), Paraná (PR), Pernambuco (PE), Piauí (PI), Rio de Janeiro (RJ), Rio Grande do Norte (RN), Rio Grande do Sul (RS), Rondônia (RO), Roraima (RR), Santa Catarina (SC), São Paulo (SP), Sergipe (SE) and Tocantins (TO).

Figure A4: Stationary Fuel Pre-tax Price Differentials in Logarithms



Note: The grey scale indicates the proportions of stationary price differential relationships for each state, based on the ADF unit root test at the 5% significance level. State abbreviations are: Acre (AC), Alagoas (AL), Amapá (AP), Amazonas (AM), Bahia (BA), Ceará (CE), Distrito Federal (DF), Espírito Santo (ES), Goiás (GO), Maranhão (MA), Mato Grosso (MT), Mato Grosso do Sul (MS), Minas Gerais (MG), Pará (PA), Paraíba (PB), Paraná (PR), Pernambuco (PE), Piauí (PI), Rio de Janeiro (RJ), Rio Grande do Norte (RN), Rio Grande do Sul (RS), Rondônia (RO), Roraima (RR), Santa Catarina (SC), São Paulo (SP), Sergipe (SE) and Tocantins (TO).

Figure A5: Stationary Fuel Pre-tax Price Differentials in Levels



Note: The grey scale indicates the proportions of stationary price differential relationships for each state, based on the ADF unit root test at the 5% significance level. State abbreviations are: Acre (AC), Alagoas (AL), Amapá (AP), Amazonas (AM), Bahia (BA), Ceará (CE), Distrito Federal (DF), Espírito Santo (ES), Goiás (GO), Maranhão (MA), Mato Grosso (MT), Mato Grosso do Sul (MS), Minas Gerais (MG), Pará (PA), Paraíba (PB), Paraná (PR), Pernambuco (PE), Piauí (PI), Rio de Janeiro (RJ), Rio Grande do Norte (RN), Rio Grande do Sul (RS), Rondônia (RO), Roraima (RR), Santa Catarina (SC), São Paulo (SP), Sergipe (SE) and Tocantins (TO).

ACKNOWLEDGMENTS

We benefited greatly from discussions with Monica Giulietti and Ana María Iregui. We also thank participants at the 2015 Agricultural & Applied Economics Association and Western Agricultural Economics Association Annual Meetings in San Francisco, as well as at the 2015 Energy and Finance Conference held at Cass Business School (City University London). We also acknowledge constructive comments and suggestions from three anonymous referees. The usual disclaimer applies.

REFERENCES

Abbott, A. and G. De Vita (2011). "Pairwise Convergence of District-Level House Prices in London." *Urban Studies*, 49: 721–740. http://dx.doi.org/10.1177/0042098011405690.

Alves, D.C.O and R. Bueno (2003). "Short-Run, Long-Run and Cross Elasticities of Gasoline Demand in Brazil." *Energy Economics*, 25: 191–199. http://dx.doi.org/10.1016/S0140-9883(02)00108-1.

ANP (2013). "Anuário Estatístico Brasileiro Do Petróleo, Gás Natural E Biocombustíveis." Agência Nacional do Petróleo, Gás Natural e Biocombustíveis.

Assis, C.A. and L. Rodrigues Lopes (1980). "A Ineficiência Da Política de Preços Para Conter O Consumo Dos Derivados de Petróleo." *Revista Brasileira de Economia*, 34: 417–428.

Bailey, N., G. Kapetanios and M.H. Pesaran (2015). "Exponent of Cross-Sectional Dependence: Estimation and Inference." *Journal of Applied Econometrics*, Forthcoming. http://dx.doi.org/10.1002/jae.2476.

Barros, C.P., L.A. Gil-Alana and P. Wanke (2014). "Ethanol Consumption in Brazil: Empirical Facts Based on Persistence, Seasonality and Breaks." *Biomass and Bioenergy*, 63: 313–320. http://dx.doi.org/10.1016/j.biombioe.2014.02.012.

Burnquist, H. and M. Bacchi (2002). "A Demanda Por Gasolina No Brasil: Uma Análise UtilizandoTécnicas de Co-Integração." In XL Congresso Brasileiro de Economia E Sociologia Rural. Equidade e Eficiênciana Agricultura Brasileira. Passo Fundo, RS.

Cheung, Y. and K.S. Lai (1995). "Lag Order and Critical Values of the Augmented Dickey-Fuller Test." *Journal of Business and Economic Statistics*, 13: 277–280.

Costa, E. and L. Neves (2015). "Transporte dutoviário de combustíveis no Brasil: desafios e oportunidades." *Revista Tecnologística*. Fedecombustivei.

Da Silva, A. C.R. Fóffano Vasconcelos, S. Pinto Vasconcelos, and R. de Mattos (2014). "Symmetric Transmission of Prices in the Retail Gasoline Market in Brazil." *Energy Economics*, 43: 11–21. http://dx.doi.org/10.1016/j.eneco.2014.02.002.

De Freitas, L.C. and S. Kaneko (2011a). "Ethanol Demand under the Flex-Fuel Technology Regime in Brazil." *Energy Economics*, 33: 1146–1154. http://dx.doi.org/10.1016/j.eneco.2011.03.011.

De Freitas, L.C. and S. Kaneko (2011b). "Ethanol Demand in Brazil: Regional Approach." *Energy Policy*, 39: 2289–2298. http://dx.doi.org/10.1016/j.enpol.2011.01.039.

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

Du, X. and M.A. Carriquiry (2013). "Flex-Fuel Vehicle Adoption and Dynamics of Ethanol Prices: Lessons from Brazil." *Energy Policy*, 59: 507–512. http://dx.doi.org/10.1016/j.enpol.2013.04.008.

Fackler, P.L. and B.K. Goodwin (2001). "Spatial Price Analysis." In Bruce Gardner and Gordon Rausser, eds, *Handbook of Agricultural Economics*, Volume 1. Amsterdam: Elsevier Science. pp 971–1024.

Ferreira, A.L., F.P. de Almeida Prado and J. da Silveira (2009). "Flex Cars and the Alcohol Price." *Energy Economics*, 31: 382–394. http://dx.doi.org/10.1016/j.eneco.2009.01.007.

Gately, D. and S. Streifel (1997). "The Demand for Oil Products in Developing Countries." World Bank Discussion Paper, 359. http://dx.doi.org/10.1596/0-8213-3891-9.

Goldberg, P. and F. Verboven (2005). "Market Integration and Convergence to the Law of One Price: Evidence from the European Car Market." *Journal of International Economics*, 65: 49–73. http://dx.doi.org/10.1016/j.jinteco.2003.12.002.

Goldemberg, J. (2006). "The Ethanol Program in Brazil." *Environmental Research Letters 1*, (014008). http://dx.doi.org/10.1088/1748-9326/1/1/014008.

Holmes, M.J., J. Otero, and T. Panagiotidis (2013). "On the Dynamics of Gasoline Market Integration in the United States: Evidence from a Pair-Wise Approach." *Energy Economics*, 36: 503–510. http://dx.doi.org/10.1016/j.eneco.2012.10.008.

Jank, M.S. (2011). "Biofuel Industry: A Local and Global View." Brazilian Sugarcane Industry Association. UNICA.

Jarque, C.M. and A.K. Bera (1987). "A Test for Normality of Observations and Regression Residuals." *International Statistical Review*, 55: 163–172. http://dx.doi.org/10.2307/1403192.

Lesage, J.P., and R.K. Pace (2008). "Spatial Econometric Modeling of Origin-Destination Flows" *Journal of Regional Science*, 48: 941–967. http://dx.doi.org/10.1111/j.1467-9787.2008.00573.x.

Leybourne, S. (1995). "Testing for Unit Roots Using Forward and Reverse Dickey-Fuller Regressions." *Oxford Bulletin of Economics and Statistics*, 57: 559–571. http://dx.doi.org/10.1111/j.1468-0084.1995.tb00040.x.

Moran, P.A.P. (1950). "A Test for the Serial Independence of Residuals." *Biometrika*, 37: 178–181. http://dx.doi.org/10.1093/biomet/37.1-2.178.

Otero, J. and J. Smith (2012). "Response Surface Models for the Leybourne Unit Root Tests and Lag Order Dependence." *Computational Statistics*, 27: 473–486. http://dx.doi.org/10.1007/s00180-011-0268-y.

Pacini, H. and S. Silveira (2011). "Consumer Choice between Ethanol and Gasoline: Lessons from Brazil and Sweden." *Energy Policy*, 39: 6936–6942. http://dx.doi.org/10.1016/j.enpol.2010.09.024.

Paul, R., D. Miljkovic and V. Ipe (2001). "Market Integration in U.S. Gasoline Markets." *Applied Economics*, 33: 1335–1340. http://dx.doi.org/10.1080/00036840010004545.

PECEGE (2009). "Custos de Produção de Cana-De-Açúcar, Açúcar e Etanol no Brasil: Safra 2008/2009." Programa de Educação Continuada em Economia e Gestão de Empresas. ESALQ/USP.

Pesaran, M.H. (2007). "A Pair-Wise Approach to Testing for Output and Growth Convergence." *Journal of Econometrics*, 138: 312–355. http://dx.doi.org/10.1016/j.jeconom.2006.05.024.

Ramsey, J.B. (1969). "Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis." *Journal of the Royal Statistical Society. Series B (Methodological)*, 31: 350–371.

Rogat, J. and T. Sterner (1998). "The Determinants of Gasoline Demand in Some Latin American Countries." *International Journal of Global Energy Issues*, 11: 162–170.

Roppa, B. (2005). "Evolução Do Consumo de Gasolina No Brasil E Suas Elasticidades: 1973 a 2003." Universidade Federal do Rio de Janeiro, Rio de Janeiro.

Rosillo-Calle, F. and L.A.B Cortez (1998). "Towards PROALCOOL II - A Review of the Brazilian Bioethanol Programme." *Biomass and Bioenergy*, 14: 115–124. http://dx.doi.org/10.1016/S0961-9534(97)10020-4.

Salvo, A. and C. Huse (2011). "Is Arbitrage Tying the Price of Ethanol to that of Gasoline? Evidence from the Uptake of Flexible-Fuel Technology." *The Energy Journal*, 32: 119–148. http://dx.doi.org/10.5547/ISSN0195-6574-EJ-Vol32-No3-5.

Salvo, A. and C. Huse (2013). "Build It, but Will They Come? Evidence from Consumer Choice between Gasoline and Sugarcane Ethanol." *Journal of Environmental Economics and Management*, 66: 251–279. http://dx.doi.org/10.1016/j.jeem.2013.04.001.

Santos, G.F. (2013). "Fuel Demand in Brazil in a Dynamic Panel Data Approach." *Energy Economics*, 36: 229–240. http://dx.doi.org/10.1016/j.eneco.2012.08.012.

Schünemann, L. (2007). "A Demanda de Gasolina Automotiva No Brasil: O Impacto Nas Elasticidades de Curto E Longo Prazo Da Expansão Do GNV E Dos Carros Flex." Faculdade de Economia e Finanças IBMEC.

Serra, T., D. Zilberman and J. Gil. (2011). "Price Volatility in Ethanol Markets." *European Review of Agricultural Economics*, 38: 259–280. http://dx.doi.org/10.1093/erae/jbq046.

Silva, G., G. Tiryaki, and L. Pontes (2009). "The Impact of a Growing Ethanol Market on the Demand Elasticity for Gasoline in Brazil." In 32nd Annual International Association for Energy Economics Conference, San Francisco. San Francisco.

Spiller, P.T. and C.J. Huang (1986). "On the Extent of the Market: Wholesale Gasoline in the Northeastern United States." *Journal of Industrial Economics*, 35: 131–145. http://dx.doi.org/10.2307/2098354.

Stigler, G.J. and R.A. Sherwin (1985). "The Extent of the Market." *Journal of Law and Economics*, 28: 555–585. http://dx.doi.org/10.1086/467101.

Suvankulov, F., M. Chi Keung Lau and F. Ogucu (2012). "Price Regulation and Relative Price Convergence: Evidence from the Retail Gasoline Market in Canada." *Energy Policy*, 40: 325–334. http://dx.doi.org/10.1016/j.enpol.2011.10.015.

Tello-Gamarra, J. (2009). "Transmissao de Precos Entre Os Mercados Do Etanol E Da Gasolina Desde O Lancamento Dos Carros Flex-Fuel, No Mercado Brasileiro." Universidade Federal do Rio Grande do Sul. Porto Alegre, Brazil.

Valdes, C. (2011). "Brazil's Ethanol Industry: Looking Forward" A Report from the Economic Research Service, BIO-02: 1–46.

White, H. (1980). "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica*, 48: 817–838. http://dx.doi.org/10.2307/1912934.

Yazgan, M. and H. Yilmazkuday (2011). "Price-Level Convergence: New Evidence from U.S. Cities." *Economics Letters*, 110: 76–78. http://dx.doi.org/10.1016/j.econlet.2010.10.013.