NEREUS

Núcleo de Economia Regional e Urbana
da Universidade de São Paulo
The University of São Paulo
Regional and Urban Economics Lab

SPATIAL PANEL DATA MODELS AND FUEL DEMAND IN BRAZIL

Gervásio F. dos Santos
Weslem R. Faria

TD Nereus 10-2012
São Paulo
2012

Spatial Panel Data Models and Fuel Demand in Brazil

Gervásio F. dos Santos and Weslem R. Faria

Abstract. The objective of this paper is to estimate the price and income elasticity for gasoline and ethanol in the fuel market for light vehicles in Brazil using spatial panel data models. Besides diversification, there are spatial features of fuel demand and supply that might influence heterogeneity in the behavior of regional consumers. Consumer behavior is heterogeneous and directly and indirectly influenced by different supply gaps and transportation and distribution costs. Thus, spatial panel data models are used to estimate price and income elasticities of gasoline and ethanol. The spatial heterogeneity and autocorrelation in the regional consumer preferences are controlled. Results showed that, aside strong evidence of the influence of spatial autocorrelation in the consumption of gasoline and ethanol, there also is a considerable competition between these fuels, mainly due to the flex-fuel technology.

1. Introduction

There is a considerable amount of literature on the estimation of fuel demand equations in the Energy Economics literature. Dahl and Sterner (1991a and 1991b) and Dahl (1995) summarized a set of principles, models and data requirements used for the estimation of the demand for gasoline and transportation fuels. Others, like Eltonny and Al-Mutairi (1993; 1995), Bentzen (1994), Espey (1996a; 1996b, 1998), Ramanathan (1999), Graham and Glaister (2002), Polemis (2006) also provided good insights on this subject. The estimation of fuel demand through panel data models can be found in Baltagi and Griffin (1983), Rouwendal (1996), Puller and Greening (1999), and Santos (2010). For the Brazilian fuel market, Burnquist and Bacchi (2002), Alves and Bueno (2003) and Roppa (2005) used time series. General results for Brazil showed that fuel demand is inelastic and that these fuels are imperfect substitutes. In general, the literature aforementioned shows that the Brazilian fuel market is very particular in its diversification and spatial heterogeneity of fuel supply. Thus, this paper aims to estimate the price and income elasticity for gasoline and ethanol in Brazil using a spatial panel data model.

In addition to the concentration in the estimation of gasoline demand, international literature presents a considerable number of papers which use time series models. Dahl and Sterner (1991a; 1991b) presented standard models to estimate price and income elasticities for gasoline mainly using monthly, quarterly and yearly time series. Eltonny and Al-Mutairi (1993; 1995) estimated the demand for gasoline in Canada and Kuwait, respectively, using cointegration techniques. Bentzen (1994) estimated the demand for gasoline in Denmark using that same technique. The same was done by Ramanathan (1999) to estimate short and long-run elasticities of gasoline demand in India. Dahl (1995) also presented a survey of demand elasticities and their components regarding the demand for transportation fuels. Espey (1996a) explained the variation in elasticity estimates of gasoline demand in the United States through meta-analysis. Espey (1996b) analyzed fuel consumption through an international automobile fuel-saving model. Espey (1998) wrote a review on gasoline demand through an international meta-analysis of elasticities. Graham and Glaister (2002) analyzed various international researches on the responses of conductors to fuel price changes; particularly the magnitude of the relevant income and price effects. This paper highlights some new results and directions in relevant literature. More recently, Polemis (2006) presented the determinants of road energy demand in Greece using cointegration techniques and vector autoregression (VAR) analysis. The evidence presented shows important differences between long-and short-run price elasticities.

In Brazil, the use of time series models with cointegration technique to estimate the price and income elasticities prevails over other models. Burnquist and Bacchi (2002) estimated demand equations for gasoline in Brazil using yearly time series. The main finding was that fuel consumption is more sensitive to income than price in both short and long run. Alves and Bueno (2003) also estimated a demand equation for gasoline using yearly time series and found that ethanol is an imperfect substitute for gasoline even in the long run. Roppa (2005) compared the competitiveness of gasoline to that of ethanol, also using yearly time series. Results showed that ethanol is an imperfect substitute for gasoline, whose consumers are indifferent to price increases in both the short and the long run. More recently, Salvo and Huse (2010) evaluated whether arbitrage tied the price of ethanol to that of gasoline, and results indicated that ethanol price increases were being driven by very high sugar prices in the international market.

Previous studies present some bottlenecks which the present study was designed to overcome. Those studies do not consider neither the substitution among the three fuels nor the use of econometric tools other than time series. Due to the heterogeneity of Brazilian economy, country-level estimations from aggregated time series might affect the results. A structure of panel data estimation might contribute to improving these kinds of studies for Brazil. One of the first studies devoted to the use of panel data to estimate energy demand equations was published by Balestra and Nerlove (1966). The authors estimated the demand for natural gas in the United States. Baltagi and Griffin (1983) estimated the demand for gasoline using panel data for OECD countries. Rouwendal (1996) analyzed the short-run behavioral responses to fuel price increases using individual consumer data on fuel usage per kilometer driven. Puller and Greening (1999) examined the household adjustment to changes in the real price of gasoline using a panel of US households over nine years.

The Brazilian fuel market for light vehicles is considerably different from other markets due to the diversity of fuels, alternative and competing. There are four main types of fuels in this market: gasoline, ethanol, vehicular natural gas (VNG) in a very small scale, and diesel. Gasoline still remains as the main fuel, but it strongly competes with ethanol. The features of the Brazilian motor vehicle fleet result in diesel not competing with other fuels. Ethanol has a historic role in the national energy policy, being an important alternative in periods of high oil prices or to help facing environmental issues. Likewise, VNG has recently been introduced in the market through a set of subsidies. Supply gaps make it hard for it to compete with ethanol and gasoline. In addition to fuel diversification, new market rules and technological advances in the automobile industry such as *flex-fuel* engines, are increasing the competition in the fuel market for light vehicles in Brazil. Furthermore, spatial features of fuel supply, such as the concentration of ethanol and gasoline production in some states might influence the heterogeneity in the behavior of regional consumers.

The main question posed by this paper is: what is the role of individual and spatial heterogeneity and in the estimation of price and income elasticities of gasoline and ethanol demand in Brazil? Thereby, the paper has a twofold. First, we estimate the price and income elasticities of gasoline and ethanol for Brazil. Second, we introduce spatial panel data models to estimate the two demand equations that control individual and spatial heterogeneity and spatial dependence of fuel consumption among Brazilian states.

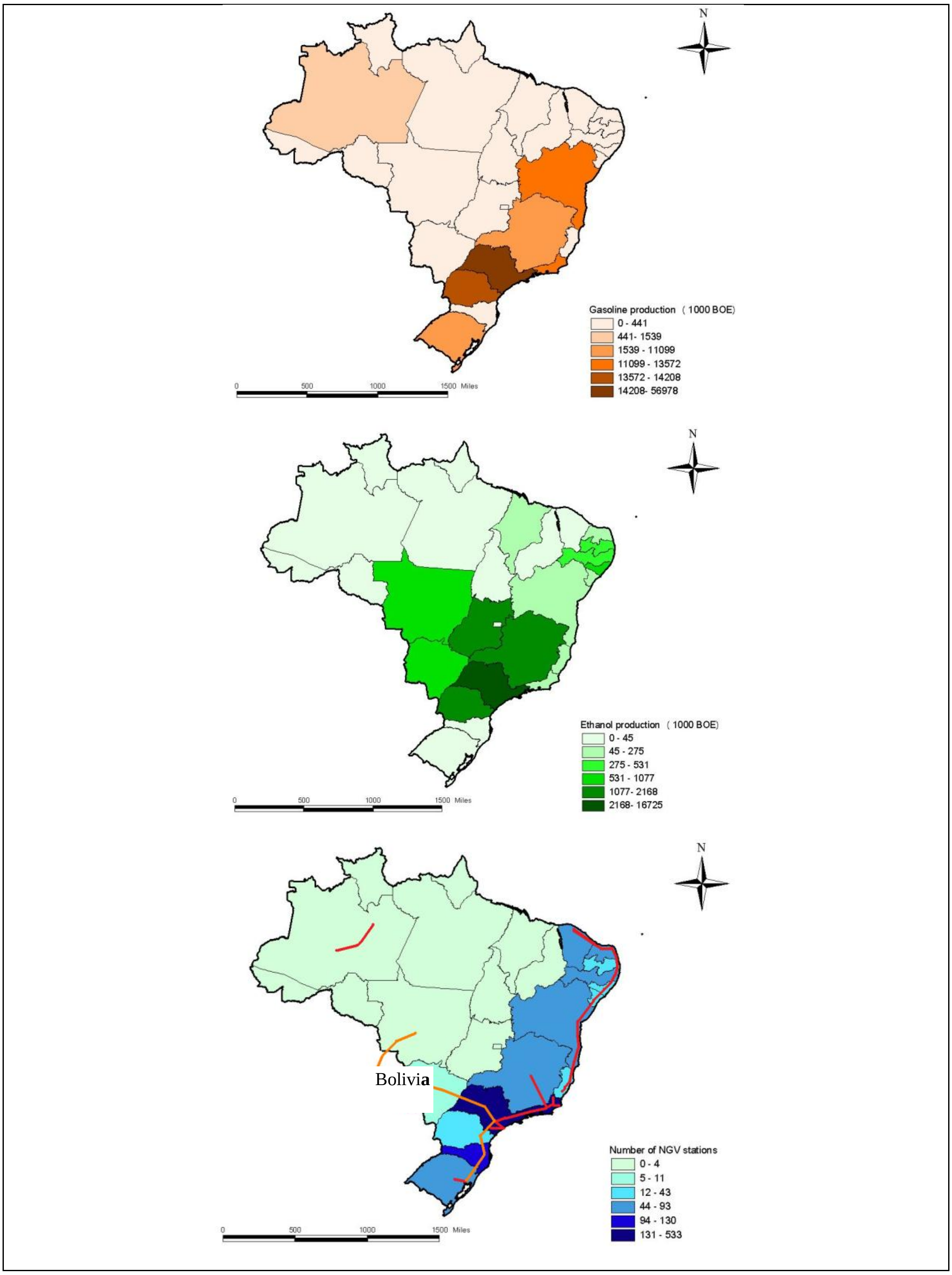
Based on the classical motivations of the regional science, spatial elements of the regional economic activity will be introduced in the fuel demand analysis. Regional scientists have formalized theory and quantitative models that rely on the notions of spatial interaction, spatial diffusion effects, hierarchies of place and spatial spillovers. In our analysis three main spatial elements need to be considered: spatial heterogeneity, spatial autocorrelation in the fuel demand and spatial autocorrelation in the idiosyncratic determinants of the fuel demand. To introduce these elements in a modeling structure spatial panel data econometric models will be used.

Past its introduction, this paper is divided in five additional sections. Section Two presents the main structural and spatial features of the Brazilian fuel market. Section Three describes econometric specifications to estimate demand equations for ethanol and gasoline. Section Four presents data requirements and equations to be estimated. Results are then presented in Section Five. Finally, Section Six presents some final remarks.

2. Structural and Spatial Features of the Brazilian Fuel Market

The dynamics of the Brazilian fuel market is still subject to the effects of its liberalization process started in 1997, fuel diversification, and technological advances in the automobile industry such as the introduction of *flex-fuel* engines. Through the well-known Law of Petroleum (Law 9.478/97) and new designs for the energy policy, the competition was introduced in the fuel market through the liberalization of price-setting of the entry of new agents. The impact this policy had on a market such as the Brazilian, with considerable diversity of substitute and complementary fuels, has drastically changed the structure of fuel demand in this country. The consumption of fuels in the road vehicle segment in Brazil amounted 63.8 million Tons of Oil Equivalent (TOE) in 2010. From this total, diesel accounts for 51.8%, gasoline for 27.0%, ethanol for 18.8%, and NGV for 2.8%. Diesel does not compete directly with other fuels, since it is mostly used by logistics trucks, agricultural machinery and some light commercial vehicles that run on diesel engines. The recent dynamics of the Brazilian fuel market are centered on ethanol demand, mainly after the introduction of the *flex-fuel* engine technology.

Figure 1. Gasoline and Ethanol Production, and Natural Gas Stations and Pipelines in Brazil, 2010

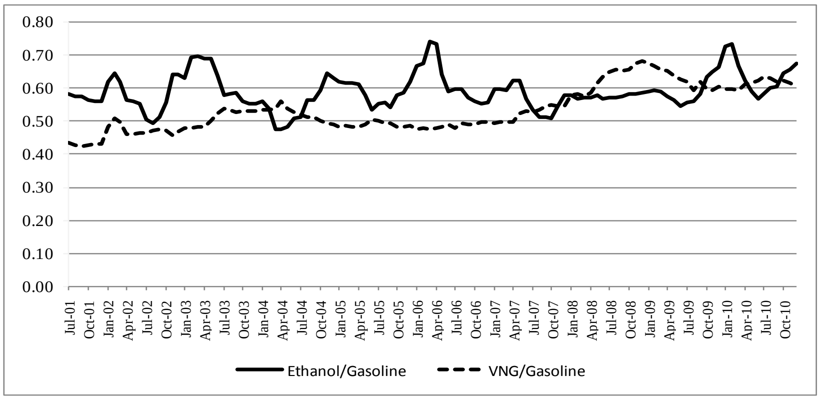


Source: Authors' elaboration based data from Brazilian Sugarcane Industry Association, National Agency of Oil and Natural Gas, and Gasnet, using softwares GeoDa and ArcView GIS 3.2.

In addition, three spatial features of fuel supply contribute in making this market very unique. First, the production of sugar-cane ethanol is concentrated in some states. The South-Central region (South and Southeast) concentrates 91.2% of the national production (Figure 1 – left) and 79.1% of ethanol consumption. Second, along with the small amount of gasoline imports in periods of high ethanol prices, its production also is concentrated in eight Brazilian states and 78.9% of this production is spatially concentrated in the South-Central region (Figure 1 – right). On the other hand, this same region concentrates 66.6% of the consumption. Finally, production and logistical problems determine that the NGV is supplied to consumers through the NGV stations only in a small portion of the national territory (Figure 1 – center). Gas pipelines operating in Brazil cover the South-Central region, which is internationally connected to the Bolivia-Brazil gas pipeline, the coastal cities of the Northeast, and a small stretch in the State of Amazonas in the North region, as shown in Figure 1. For this reason, competition is only effective for two of these fuels: gasoline and ethanol. Moreover, there might be different patterns of substitution between these two fuels that determine a considerable spatial heterogeneity in regional consumer behavior. Taking these elements into account when estimating demand equations can provide important insights for the literature and also for future energy and regulatory policymaking in Brazil.

Diversification of the fuel supply in Brazil stems from the national energy policy to increase the substitution of gasoline. Ethanol was introduced in the Brazilian energy matrix in 1975 through the *Pro Alcool*, the National Ethanol Fuel Program. This was a large scale energy production program designed to substitute vehicular fossil fuels by sugar-cane ethanol and created due to the first and second oil shocks of 1974 and 1979, respectively. Likewise, the decision to produce ethanol from sugarcane was driven by the low sugar prices in the international commodity market at that time. Since the beginning of *Pro Alcool*, ethanol fuel was massively introduced as a complementary and substitute fuel in Brazil. Initially, ethanol was mixed into gasoline given its features of complementary good. Nowadays, this mixing is still done, at a rate of around 20% to 25% of ethanol mixed in the gasoline. Additionally, the Brazilian automobile industry started producing vehicles with engines that ran solely on ethanol. For this reason, ethanol also figures as a substitute for gasoline. The program was effective until the first half of the 1980's. After that, a set of factors such as the decline of oil prices (and consequently of gasoline) along with increases in sugar prices in international markets, led to the termination of the program. In addition, the engagement of Brazil in liberal policies which enforced the elimination of subsidies made the program unfeasible. In spite of that, the production of ethanol and vehicles with ethanol engines continued in small scale.

Figure 2. Price Relations in the Brazilian Fuel Market – (Jul/2001-Dec/2010)



Source: ANP – Brazilian National Agency of Oil and Biofuel.

Regarding energy efficiency, ethanol's calorific value of ethanol is lower than that of gasoline. Therefore, technical limitations of engines make ethanol less efficient that gasoline. As a consequence, the price of ethanol should remain around 70% of the price of gasoline in order for it to maintain competitiveness in the Brazilian fuel market. In 1989, when the *Pro Alcool* collapsed, this percentage was larger than 75% (Roppa, 2005). Since then, this price relation has been rigorously maintained (Figure 02) mainly due to government assistance through measures such as setting lower fuel taxes[1]. The end of *Pro Alcool* in 1989 did not eliminate the use of ethanol as fuel. A considerable fleet of vehicles running on ethanol still remained in the fuel market. Sales of this type of vehicles, nonetheless, declined from 80% of total sales in the period of 1983-1987 to 0.7% in 2000. Data from the National Department of Transport (2008) indicate that in 2006 the national motor vehicle fleet was composed by 45.6 million vehicles. About

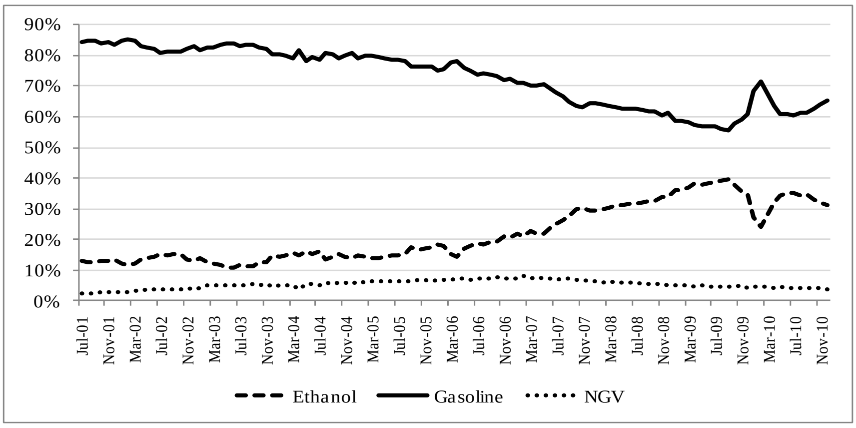
[1] In order to subsidize ethanol and NGV, the tax named *Contribution of Economic Domain* (CIDE) has not been charged on these two fuels.

22.0 million were light vehicles; and, from this total, 3.0 million were pumped exclusively with ethanol, which represented 13.6% of the light vehicle fleet. Maintaining these ethanol-consuming vehicles was important to foster new developments in the technology of large scale ethanol production and also for its very existence.

The recent increase in the demand for ethanol in Brazil results from a revolution in the automobile industry: the development of *flex-fuel engines*[2]. This technology allows pumping the same vehicle with either gasoline or ethanol. It was developed in the United States in the 1980's and has been used in Brazil since the 1990's; in the end of 2003, it was introduced in the market; in 2005, sales of *flex-fuel* vehicles was greater than that of vehicles running solely on gasoline; in 2010, sales of cars and light vehicles amounted 3.51 million in Brazil. From this total, 81.8% (2.87 million) had *flex-fuel* engines. In the meantime, VNG was introduced in the fuel market. The initial strategy was to use it as a substitute for diesel in large road vehicles. However, due to logistical problems, its usage was restricted to large urban centers. In 1994, subsidies allowed for the emergence of some VNG demand pools in Brazil, mainly to be used in urban transportation such as city buses and taxis. However, this was not sufficient for this fuel to thrive. As shown in Figure 3, gasoline still remains as Brazil's primary fuel, despite the fact that its market share declined from 84.4% in 2001 to 65.2% in the end of 2010. In that same period, both ethanol and VNG had an increase in consumption. The first going from 13.2% to 31.2% and the latter from 2.5% to 3.5%.

[2] This technology is embedded in the *Poli-fuel* engines trend of the world automobile industry.

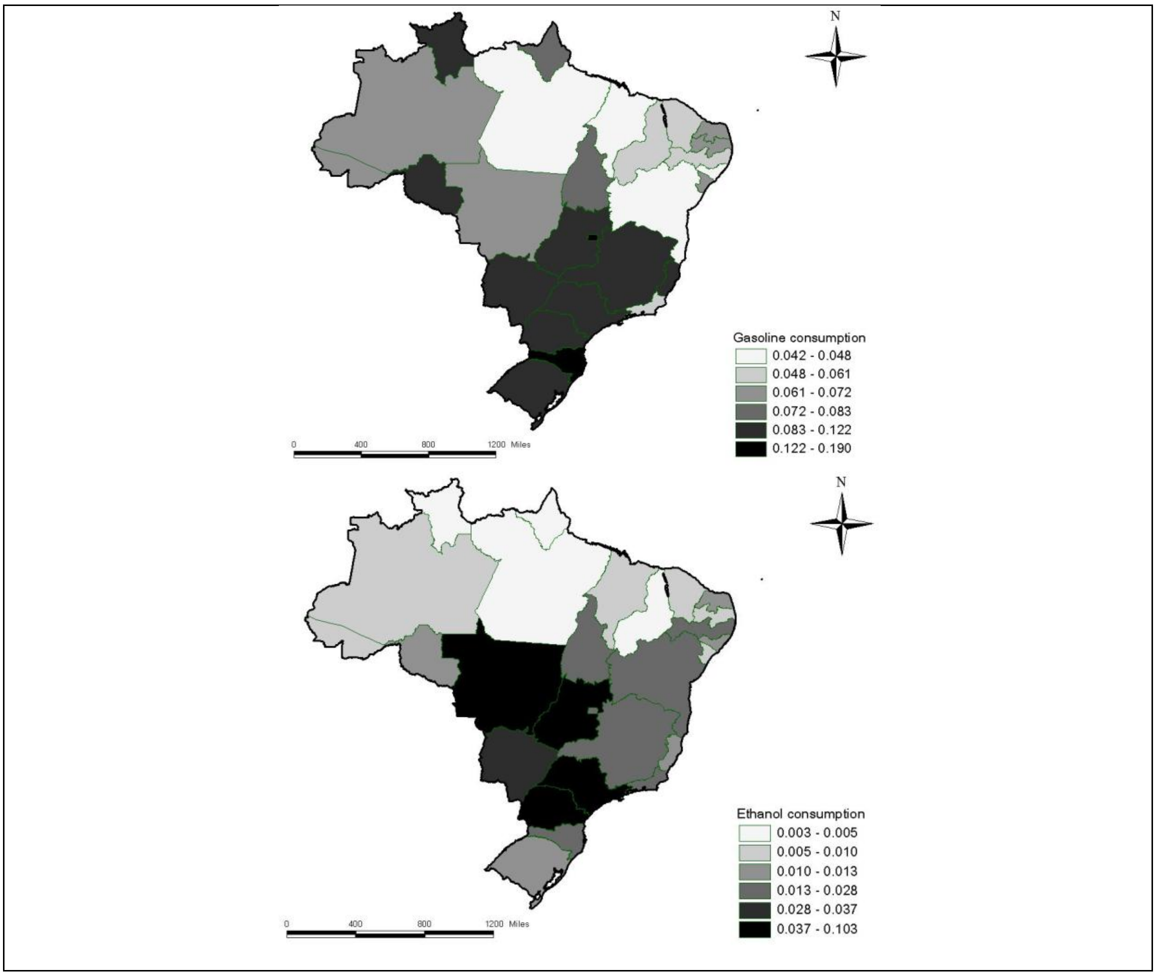
Figure 3. Market Share of Gasoline, Ethanol and NGV in Brazil (Jul/2001-Dec/2010)



Source: ANP – National Agency of Oil and Biofuel.

Taking the data aforementioned into account, this paper searches to present different supply structures that affect the behavior of the regional fuel consumer. For this reason, we propose to control spatial autocorrelation and individual heterogeneity in the estimation of gasoline and ethanol demand equations using spatial panel data models. Price and consumption data for these two fuels are collected in the whole country. Figure 4 shows the per capita consumption of gasoline and ethanol in 2010 in Brazil. The consumption of both fuels was divided into six groups, ranging from lowest to largest per capita consumption. The highest consumption of gasoline was observed in Santa Catarina and Distrito Federal, followed by states in the Southeast (except for Rio de Janeiro) and in the South. On the other hand, the lowest consumption was that of the North and Northeast regions. High gasoline consumption is spatially concentrated in the richest states, which reflects its use in large and luxurious vehicles. Regarding ethanol, although the highest consumption was verified in the two richest states in the country - Sao Paulo and Paraná - two other states - Goiás and Mato Grosso - also ranked among the highest in consumption of ethanol. Its per capita consumption is not as spatially concentrated as that of gasoline, what reflects the scattering of the consumption mainly due to its use in smaller and more “popular” vehicles in Brazil.

Figure 4. Brazilian per capita consumption of gasoline and ethanol in 2010



Source: Authors' elaboration based on softwares GeoDa and ArcView GIS 3.2.

A part from the influence of the supply and demand heterogeneity in the estimation of demand equations for gasoline and ethanol, other spatial elements need to be considered. We based on the classical motivations to expect that the spatial dependence of regional economic activity also influences the fuel demand. The analysis of these classical motivations is presented by LeSage (1999). The regional science field is based on the premise that location and distance are important forces in the determination of economic activity. Thus, regional scientists have formalized theory and quantitative models that rely on the notions of spatial interaction, spatial diffusion effects, hierarchies of place and spatial spillovers.

In our analysis there are three main spatial elements need to be considered: spatial heterogeneity, spatial autocorrelation in the fuel demand and spatial autocorrelation in the idiosyncratic determinants of the fuel demand. First, spatial heterogeneity refers to variation in relationships over space. By hypothesis, there might be a different relationship for the fuel demand and its determinants for every point in space. As a consequence, individual heterogeneity of the spatial units needs to be controlled in the econometric estimation.

Second, we consider that for several reasons consumer preferences might be spatial correlated in a way that the fuel consumption in a spatial unit affects the fuel consumption in other units. This requires the econometric modeling of the spatial lag of the fuel consumption. Finally, the intensity of the fuel consumption in a spatial unit can stimulate responses in the supply side in this unit and in the neighbors units. The idiosyncratic factors that determine the fuel consumption also are spatially correlated. This hypothesis requires the econometric modeling of the spatial correlation in the error term. To introduce the spatial elements above in the modeling will be necessary to handle the spatial panel data models, which will be presented in the next section.

3. Methodology

3.1. Spatial Models and Tests

Usage of spatial econometrics is widespread in analysis of cross-section models. The main reason to incorporate spatial components in the models is to control for spatial dependence, which can be introduced into linear regression models in two ways. Firstly, by introducing spatial dependence through the creation of new dependent and independent variables and of spatial dependence in the error term in the model. These new variables capture spatial dependence as a weighted average of the neighboring values. Models that use this procedure are classified as spatial lag models. Secondly, by introducing a spatial autoregressive error process (or moving averages). These kinds of models are classified as spatial error models (Anselin, 1992; 1999a; 1999b). The failure to introduce spatial dependence in the form of spatial lag into the model can make the OLS estimators biased and inconsistent, regardless of the behavior of the error terms. As for spatial error terms, the consequence of their omission is the inefficiency of the OLS estimators (Anselin, 1988; Anselin and Bera, 1998).

The application of spatial analysis to panel data is still quite restricted, mainly due to two reasons: the theoretical models are very recent and the computation implementation is still limited. Spatial panel theoretical models, as well as the investigation of its asymptotic properties and specification tests, are new to econometric literature. The studies differ from each other basically in the procedure used to calculate the estimators, usually based on the Maximum Likelihood Function (MLF) taking into account the fixed effect models (Elhorst, 2005; Yu *et al.*, 2008; Lee and Yu, 2010a; Lee and Yu, 2010b). Others have considered both fixed and random effect models (Elhorst, 2003, Anselin *et al.*, 2008). In addition, others are based on the Generalized Moments (GM) estimators (Kapoor *et al*., 2007; Elhorst *et al.*, 2010). Aside from being new in the literature, there is no available program to directly perform the estimations of these models. However, some computational routines might be adapted, such as Elhorst's and Prucha's routines for Matlab and Stata, respectively.

Specification tests to assist the process of choosing the most suitable model represent another important step in the development of spatial panel data models. Most of the tests developed so far are devoted to verifying the existence of spatial correlation in random effect models. Baltagi *et al*. (2003) developed Lagrange Multiplier (LM) tests to verify the joint significance of the presence of regional random effects and spatial correlation on the error term and also the existence of spatial autocorrelation on the error term assuming regional random effects. Then, Baltagi *et al.* (2007b) introduced serial correlation in those tests, which allowed for the simultaneous assessment of spatial and serial correlation. Baltagi *et al.* (2008) also developed an LM test to verify the individual presence of spatial autocorrelation due to spatial lag dependence and another set of tests to verify the presence of spatial lag dependence on random effects models. Finally, Baltagi *et al*. (2007a) presented unit root tests taking spatial dependence into account.

In this study, we estimate two demand equations - for gasoline and ethanol – in order to obtain the price, cross-price and income-elasticities of demand for both these fuels considering the broader range of models available for simultaneous analysis of panel data and spatial dependence. We start from a basic demand equation:

$$y_{it} = f\left(p_{it}^{g}, p_{it}^{s}, I_{it}\right) \tag{1}$$

y_{it} is the per capita consumption of gasoline (or ethanol), $p^{g}{}_{it}$ is the real price of the respective fuel, $p^{s}{}_{it}$ are the prices of substitute fuels, and I_{it} the real per capita income. This model is easy to estimate, to interpret and does not over-demand data requirements, see Dahl and Sterner (1991). For all variables in a panel data structure the index *i* represents the 27 Brazilians states while the index *t* represents quarterly periods of time.

Before describing the taxonomy of spatial models, consider the standard fixed effects model, in which *i* is the cross-section unit index and *t* is the time index (Baltagi, 2001):

$$y_{it} = \alpha_i + x_{it}\beta + \varepsilon_{it} \tag{2}$$

where $i = 1, \dots, N$; $t = 1, \dots, T$; y_{it} is the dependent variable; x_{it} is a vector of explanatory variables; α_i are the time-invariant individual components and $\sum_i \mu_i = 0$, in order for that component to be separately identifiable from the constant term; and ε_{it} is the error term. The vector β represents the parameter(s) to be estimated. Two equations will be estimated.

The fixed effects spatial lag model in stacked form can be described as (Elhorst, 2003; Anselin *et al*., 2008):

$$y = \rho(I_T \otimes W_N)y + (i_T \otimes \alpha) + X\beta + \varepsilon \tag{3}$$

where y is an $NT \times 1$ vector; X is a $NT \times K$ matrix; ε is a $NT \times 1$ error term vector; ρ is the spatial autoregressive parameter; β is a $K \times 1$ vector of parameter(s) to be estimated; α is an $N \times 1$ vector of individual fixed effects, with the analogue constraint $\mu' i_N = 0$; W is an $N \times N$ positive non-stochastic spatial weight matrix; I_T is an identity matrix of dimension T; and $E[\varepsilon\varepsilon'] = \sigma^2 I_{NT}$ and $E[\varepsilon] = 0$. Due to the stacking of cross-sections, this approach differs from the classic fixed effects analysis in the formulation of the Kronecker product.

The fixed effects spatial error model can be written as (Elhorst, 2003; Anselin *et al.*, 2008):

$$y = (i_T \otimes \alpha) + X\beta + u \tag{4}$$

and

$$u = \lambda(i_T \otimes W_N)u + \varepsilon \tag{5}$$

where y, X, β, α, and W are as described above; λ is the spatial autoregressive coefficient; ε is an $N \times 1$ idiosyncratic error vector; and u is an $N \times 1$ vector of spatial autoregressive (SAR) error term.

The analysis of fixed effects lag and spatial error models is similar to that of cross-section spatial models. Moreover, like the analysis of cross-section, both fixed effect models discussed in this work, spatial lag and spatial error will be estimated by maximum likelihood (ML). The derivation of Maximum Likelihood Estimators (MLE) is presented and discussed by Elhorst (2003; 2005) and Anselin *et al.* (2008). The estimators of random effects models are also ML, but estimation by GM is presented for such models.

The traditional model of random effects can be written as follows (Baltagi, 2001):

$$y_{it} = x_{it}\beta + \varepsilon_{it} \tag{6}$$

where:

$$\varepsilon_{it} = \mu_i + v_{it} \tag{7}$$

where $i = 1, \ldots, N$; $t = 1, \ldots, T$; $\mu_i \sim IID(0, \sigma_u^2)$ is the cross-sectional random component; $v_{it} \sim IID(0, \sigma_v^2)$ is the idiosyncratic error term; and μ_i and v_{it} are independent. The equation (6) can be rewritten for $t = 1, \ldots, T$:

$$\varepsilon_t = \mu + v_t \qquad (8)$$

where ε_t and v_t are $N \times 1$ error vectors; and μ is an $N \times 1$ vector of cross-sectional random components.

As noted by Anselin *et al.* (2008), the incorporation of spatial components in a random effects model is basically done through the inclusion of spatial correlation of the error in the error term of regression. Thus, the point of analysis is focused on the random effects spatial error model, since the random effects spatial lag model can be obtained more directly by correctly specifying the structure of the variance-covariance error matrix. In order to incorporate the spatial error term, it is assumed that the error term of equation (7) follows an autoregressive process like (Baltagi *et al.*, 2007b):

$$v_t = \lambda W_N v_t + u_t \qquad (9)$$

where $t = 1, \ldots, T$; λ is the scalar spatial autoregressive coefficient with $|\lambda| < 1$; and u_t is an *i.i.d* idiosyncratic error term with a σ_u^2 variance . Equation (8) can be rewritten as:

$$v_t = (I_N - \lambda W_N)^{-1} u_t = B_N^{-1} u_t \qquad (10)$$

where W_N satisfies the condition that $(I_N - \lambda W_N)$ is non-singular for all λ; I_N is an identity matrix of dimension N; and $B_N = (I_N - \lambda W_N)$. Then, the equation (7) can be written in stacked form as:

$$\varepsilon = (i_T \otimes I_N)\mu + (I_T \otimes B_N^{-1})u \qquad (11)$$

where ε is the $NT \times 1$ vector of error terms; and $u \sim IID(0, \sigma_u^2 I_{NT})$ is the $NT \times 1$ vector of idiosyncratic errors. The variance–covariance matrix of ε is (Anselin *et al.*, 2008):

$$E[\varepsilon\varepsilon'] = \sigma_\mu^2(i_T i_T' \otimes I_N) + \sigma_u^2(I_T \otimes (B_N' B_N)^{-1}) \qquad (12)$$

The first term on the right side of equation (11) specifies the serial correlation (correlation in the time dimension) without considering the spatial correlation (correlation in the cross-section dimension), while the second term specifies the spatial correlation, but not the serial correlation. All models estimated in this work, fixed and random effects, consider the presence of serial correlation. Fixed and random effects are considered due to the unobserved heterogeneity in the per capita consumption of fuel in the Brazilian states. This heterogeneity derives from elements such as the different regional supply conditions, different per capita income levels and others, as described in the previous sections. Also, the use of a panel data set with 38 time periods, as described later in this section, makes it important to consider the presence of serial autocorrelation.

Baltagi *et al*. (2003) also define two main types of LM tests to verify the existence of regional random effects and spatial correlation on the error term, jointly and individually. This study uses these tests as criteria for identifying the most appropriate model estimated. The purpose of the application of these tests is to obtain additional information, since the relevant literature still lacks more specific tests for spatial panel analysis, for example, under the assumption of fixed effects.

The Joint LM test is a One-Sided Joint Test and it tests for the simultaneous presence of regional random effects and spatial correlation on the error term. Thus, considering the notation used in the previous description of the random effects model, the null hypothesis is $H_0\colon \sigma_u^2 = \lambda = 0$. Defining $\theta = \left(\sigma_v^2, \sigma_\mu^2, \lambda\right)'$, the LM statistic can be described as:

$$LM = \widetilde{D}_\theta' \tilde{J}_\theta^{-1} \widetilde{D}_\theta \qquad (13)$$

where $\widetilde{D}_\theta = (\partial L/\partial\theta)\left(\tilde{\theta}\right)$ is a 3×1 vector of partial derivatives with respect to each element of θ, evaluated at the restricted (MLE) $\tilde{\theta}$; and $\tilde{J} = E[-\,\partial L/\partial\theta\partial\theta']\left(\tilde{\theta}\right)$ is the information matrix corresponding to θ, evaluated at the restricted MLE $\tilde{\theta}$. According to Baltagi *et al.* (2003), under the null hypothesis, the variance–covariance matrix of the error term in the random effect model (equations (5) and (6)) is reduced to $\sigma_v^2 I_{TN}$. In other words, under the null hypothesis, using the traditional OLS panel model is adequate.

The LM individual test is a Conditional Test which tests for the presence of spatial correlation on the error term, assuming the existence of a random effects model, i.e., if $\sigma_\mu^2 > 0$. The null hypothesis is $H_0: \lambda = 0$. Under the null hypothesis, the variance–covariance matrix of the error term in the random effect model (equations (5) and (6)) is reduced to $\left(\sigma_\mu^2 J_T \otimes I_N\right) + (\sigma_v^2 I_{NT})$. This means that, under the null hypothesis, the traditional random effect panel model is more appropriate than the random effect spatial model.

3.2. Data

The demand equations to be estimated are:

$$lnG_{it} = \beta_0 + \beta_1 lnP_{G(it)} + \beta_2 lnP_{E(it)} + \beta_3 lnGDP_{it} + \epsilon_{it} \tag{14}$$

$$lnE_{it} = \beta_0 + \beta_1 lnP_{G(it)} + \beta_2 lnP_{E(it)} + \beta_3 lnGDP_{it} + \epsilon_{it} \tag{15}$$

where the variable *G* is the per capita consumption of gasoline and the variable *E* is the per capita consumption of ethanol. Regarding the controls, P^G is the price of gasoline, P^E is the real price of ethanol and GDP is the per capita Gross Domestic Product. Variables *i* and *t* represent a panel composed by quarterly data set from the 27 Brazilian states for the period from Jul/2001 to Dec/2010, the period in which the ANP collected data on fuel prices and consumption. The quarterly price index and population data were obtained from the Brazilian Institute of Geography and Statistics (IBGE). Since there are no quarterly data on the GDP of each state, in Brazil, a proxy variable had to be used. This variable was the Product and Service Trading Tax (ICMS) obtained from the Brazilian National Treasury. To avoid an endogeneity problem we used the net ICMS, obtained from specific fuel ICMS.

Regarding the non-inclusion of VNG prices in the equation, aside from the small share of 2.8% in the amount of fuel consumed in Brazil, other considerations should be made. First, only 10 of the 27 Brazilian states have a considerable presence of natural gas stations, making it impossible to construct a panel database for all the states, even one that is unbalanced, since eight states have either one or no natural gas stations. Second, in the majority of these states natural gas stations are only located just in the capital. Third, the usage of VNG still is very restricted, and is most common in urban transportation vehicles (buses and taxis) in the capitals of these states. Lastly, in this market the price is "artificially" low compared to other fuels due to this market's growth-stimulating policies. For this reason, we think that there is no real competition among VNG and other fuels in Brazil that could be captured by the demand equations proposed in this study considering a panel database for the Brazilian states.

Finally, it is noteworthy that this study uses routines developed under the Project-R for estimation of spatial models[3], and these routines are still under development and do not meet all modeling possibilities for panel data. That is due to the fact that some methodological improvements, such as the tests of specification for defining the most suitable model, are still under development and maturation.

4. Results

Global spatial autocorrelation tests were carried out for two dependent variables used in the models before the econometric estimation: per capita consumption of gasoline and per capita consumption of ethanol. The purpose of these tests was to provide additional information about the presence of spatial elements in the variables of interest. The data used refer to each state and were annualized to test for the presence of global spatial autocorrelation. Estimation results presented ahead consider data at quarterly level. These tests are Unconditional Tests, since they are performed on the dependent variables of the models to be estimated, with no consideration toward specifications or different strategies for identifying these models. Unconditional Tests were performed based on Moran's *I*, which is well known in the literature for assessing the degree of

[3] These routines correspond to the package "splm" for simultaneous analysis of panel data and spatial dependence. This package is under administration of Gianfranco Piras and Giovanni Millo and is available at http://r-forge.r-project.org/projects/splm/.

global spatial autocorrelation of the variable of interest.[4] We used two approaches as contiguity criterion: Queen and *k* nearest neighbors spatial weight matrices. The Queen spatial weight matrix represents a binary definition of neighborhood, which considers neighboring spatial units of a region. These spatial units have at least one common point. The *k* nearest neighbors weight matrix considers the distance among regions as a neighboring criteria.[5]

These two neighborhood criteria were used to verify the robustness of results, which are presented in Tables 1 and 2 for per capita gasoline consumption and per capita ethanol consumption in Brazil, respectively. Results presented in both tables indicate there is global spatial autocorrelation [at the 1% level of statistical significance for both variables in all years.[6] Global spatial autocorrelation tests conducted on these variables are unconditional and, therefore, represent only an indication about the spatial correlation that can be found in the estimations. However, other procedures were used to verify this, basically via significance analysis of the spatial components coefficients and performance of LM tests. Moran's *I* values were all positive indicating that, overall, regions with a high per capita consumption of gasoline or ethanol are surrounded by regions of high per capita consumption of gasoline or ethanol and vice-versa.

[4] Moran's *I* is formally presented as:

$$I_t = \left(\frac{n}{S_0}\right)\left(\frac{Z_t' W Z_t}{Z_t' Z_t}\right), t = 1, \ldots, n$$

where Z_t is a vector of *n* observations for the year *t* deviated from the mean for the variable of interest, i.e., deforested area. *W* represents the spatial weight matrices such that: 1) the diagonal elements W_{ii} are equal to zero and; 2) the non-diagonal elements W_{ij} indicate the way that a region *i* is spatially connected with the region *j*. *S* is a scalar term that is equal to the sum of all *W* elements. Moran's *I* provides a very good approximation of any linear association between the vectors observed at time *t*, and the weighted average of neighboring values, or spatial lags (Cliff and Ord, 1981).

[5] The weight matrix of type *k* nearest neighbors is a recommended solution when the weights in the distance and the units of area are very irregular. Further details about the *k* nearest neighbors spatial matrix can be found in Le Gallo e Ertur (2003). The matrix of *k* nearest neighbors was also used by Pace and Barry (1997), Pinkse and Slade (1998) and Baller *et al*. (2001) in different applications.

[6] A randomized procedure of Moran Indexes was used to run this test (details about the procedure can be found in Anselin, 2005).

Table 1. Moran's *I* for the Gasoline Consumption

Year	*Spatial Weight Matrix*	*Moran's I*	*Mean*	*SD*	*p-value*
	Queen	0.511	-0.038	0.134	0.001
2002	*3 nearest neighbors*	0.478	-0.039	0.129	0.001
	5 nearest neighbors	0.499	-0.036	0.098	0.001
	Queen	0.509	-0.039	0.130	0.001
2003	*3 nearest neighbors*	0.480	-0.039	0.125	0.001
	5 nearest neighbors	0.483	-0.039	0.091	0.001
	Queen	0.525	-0.041	0.129	0.001
2004	*3 nearest neighbors*	0.503	-0.037	0.137	0.002
	5 nearest neighbors	0.475	-0.040	0.096	0.001
	Queen	0.511	-0.035	0.133	0.001
2005	*3 nearest neighbors*	0.489	-0.043	0.128	0.003
	5 nearest neighbors	0.461	-0.036	0.095	0.001
	Queen	0.491	-0.042	0.127	0.001
2006	*3 nearest neighbors*	0.467	-0.038	0.125	0.001
	5 nearest neighbors	0.429	-0.043	0.091	0.001
	Queen	0.478	-0.039	0.131	0.002
2007	*3 nearest neighbors*	0.485	-0.036	0.124	0.001
	5 nearest neighbors	0.426	-0.043	0.091	0.001
	Queen	0.279	-0.045	0.132	0.009
2008	*3 nearest neighbors*	0.244	-0.043	0.128	0.010
	5 nearest neighbors	0.208	-0.038	0.095	0.010
	Queen	0.431	-0.042	0.132	0.001
2009	*3 nearest neighbors*	0.408	-0.035	0.128	0.001
	5 nearest neighbors	0.328	-0.038	0.098	0.003
	Queen	0.410	-0.036	0.132	0.005
2010	*3 nearest neighbors*	0.353	-0.043	0.127	0.004
	5 nearest neighbors	0.242	-0.037	0.093	0.009

Source: Authors' elaboration based on softwares *R*, GeoDa and ArcView GIS 3.2.

Table 2. Moran's *I* for the Ethanol Consumption

Year	*Spatial Weight Matrix*	*Moran's I*	*Mean*	*SD*	*p-value*
	Queen	0.615	-0.034	0.143	0.002
2002	*3 nearest neighbors*	0.597	-0.035	0.136	0.001
	5 nearest neighbors	0.580	-0.040	0.106	0.001
	Queen	0.640	-0.047	0.140	0.001
2003	*3 nearest neighbors*	0.605	-0.037	0.134	0.001
	5 nearest neighbors	0.563	-0.032	0.102	0.001
	Queen	0.630	-0.030	0.140	0.001
2004	*3 nearest neighbors*	0.614	-0.036	0.134	0.001
	5 nearest neighbors	0.549	-0.040	0.093	0.001
	Queen	0.604	-0.028	0.143	0.001
2005	*3 nearest neighbors*	0.597	-0.042	0.128	0.001
	5 nearest neighbors	0.562	-0.039	0.098	0.001
	Queen	0.456	-0.038	0.125	0.002
2006	*3 nearest neighbors*	0.454	-0.041	0.119	0.003
	5 nearest neighbors	0.430	-0.038	0.086	0.001
	Queen	0.421	-0.038	0.129	0.003
2007	*3 nearest neighbors*	0.405	-0.035	0.119	0.002
	5 nearest neighbors	0.404	-0.039	0.092	0.001
	Queen	0.428	-0.035	0.134	0.001
2008	*3 nearest neighbors*	0.444	-0.043	0.130	0.001
	5 nearest neighbors	0.436	-0.041	0.089	0.001
	Queen	0.438	-0.042	0.135	0.002
2009	*3 nearest neighbors*	0.462	-0.045	0.127	0.002
	5 nearest neighbors	0.457	-0.035	0.098	0.002
	Queen	0.348	-0.039	0.129	0.001
2010	*3 nearest neighbors*	0.329	-0.039	0.122	0.004
	5 nearest neighbors	0.345	-0.037	0.098	0.006

Source: Authors' elaboration based on softwares *R*, GeoDa and ArcView GIS 3.2.

From 2002 to 2005, spatial autocorrelation was stronger for ethanol consumption than for gasoline consumption, while from 2006 to 2010, it followed the opposite pattern. There is no clear explanation for this result. However, it must be noted that before 2005 ethanol consumption was restricted to vehicles that ran solely on ethanol (*non–flex-fuel)*, so the higher and lower per capital ethanol consumption might be related to some supply factor. After 2005, the *flex-fuel* fleet started increasing throughout the country, leading to the concentration of gasoline consumption in higher income states. Another feature observed in the results is that the global spatial correlation had a higher coefficient for the Queen matrix than it did for the *k* nearest neighbors matrix, in all years and for the consumption of both fuels. This result might be influence by the number of spatial units (27 Brazilian states). Thus, estimates were carried out using the Queen criterion of contiguity.

Econometric results are presented in Tables 3 and 4. Both tables were divided in three parts, according to the estimation method: pooled OLS, fixed effects and random effects. For all these cases the models were estimated considering the traditional structure of panel data and the structure of different spatial panel models, spatial lag and/or error. For pooled OLS and fixed effects, spatial lag and spatial error models were estimated separately. For random effects models, spatial lag and spatial error models were estimated separately and spatial error and lag were estimated together. The latter random effects model was estimated by both ML and GM. With the exception of the GM method, all other spatial models were estimated by ML.

Presenting the widest possible set of results of spatial models enables us to check under different hypothesis, both in terms of the way that the spatial correlation may be present (spatial lag or error) and the type of estimator, whether the estimates of elasticities differ in sign and magnitude from one model to another. This is justified by the lack of specification tests with a large capacity for evaluating models that can assist in identifying the most appropriate model to be used.

Results in Table 3 – for the estimation of price elasticity of demand for gasoline – show that the coefficients of the spatial lag and error components of pooled OLS, fixed effects and random effects models are all significant at the 1% of statistical level. Regarding spatial error models, the significant value of the *lambda* (λ) indicates that spatial effects not modeled are a part of the error in such models. Spatial lag models indicate, by mean of the *rho* (ρ) term, the spatial lag of the dependent variable, i.e. the average per capita gasoline consumption in a region's neighboring areas is important to explain its own consumption of gasoline. Moreover, spatial lag models suggest there are significant spatial spillover effects around the activity of gasoline consumption. It is worth highlighting that all spatial model estimations consider serial correlation. For the random effects model, Baltagi *et al*. (2003) rewrite the equation (12) so that it defines a term $\phi = \sigma_\mu^2/\sigma_u^2$. This measure may be important in checking the adequacy of random effect models to the analysis. Besides, it may bring additional information about the relationship between serial and spatial correlation. The ϕ values of spatial error, spatial lag and spatial lag+error random effects models estimated by ML in Table 3 were significant at a 1% statistical level. This is an indication that serial correlation should be considered, mainly for the specification of models by random effects.

Table 3 also shows the results of the estimation of the spatial random effects model by GM. Starting from the basic model $y_N = x_N\beta + u_N$, Kapoor et al. (2007) introduces spatial correlation, for $t = 1, \dots, T$, by specifying $u_N = \rho(I_T \otimes W_N) + \varepsilon_N$ and serial correlation as $\varepsilon_N = (e_T \otimes I_N)\mu + v_N$. This notation is analogous to that presented in section 3.1, with the exception of ρ, which, in this case, is the scalar autoregressive parameter; and e_T is a unit vector. The difference between this procedure and the ML estimators lies in the way that variance-covariance works to define spatial spillovers. The ϕ was created for ML estimation and, analogously, for GM estimation we have that $\sigma_1^2 = \sigma_v^2 + T\sigma_\mu^2$ and $\theta = 1 - \sigma_v^2/\sigma_1^2$. For the model in Table 3, the θ term was equal to 0.928, indicating a high composition of the variance due to specific effect (μ). Since σ_μ^2 is large compared to σ_v^2, that means the random effects may be appropriate in such situation.

Table 3. Price-Elasticity of Demand for Gasoline: Estimates for Brazil

Methods	*Models*	*Constant*	*Rho (ρ)*	*Lambda (λ)*	*Gasoline Price*	*Ethanol Price*	*GDP*
Pooled OLS	*Non-Spatial Models*	0.068***			-0.031***	0.018***	0.760***
		(0.004)			(0.004)	(0.029)	(0.005)
	Spatial Lag Model	0.607***	0.332***		-1.237***	0.514***	0.542***
		(0.028)	(0.114)		(0.117)	(0.077)	(0.021)
	Spatial Error Model	-0.094		0.358***	-1.209***	0.536***	0.629***
		(0.130)		(0.033)	(0.138)	(0.090)	(0.021)
Fixed Effect	*Non-Spatial Models*	0.046***			-0.003	0.007***	0.279***
		(0.002)			(0.002)	(0.002)	(0.021)
	Spatial Lag Model		0.368***		-0.412***	0.069	0.169***
			(0.033)		(0.106)	(0.059)	(0.035)
	Spatial Error Model			0.383***	-0.512***	0.153**	0.149***
				(0.033)	(0.106)	(0.064)	(0.035)
Random Effect	*Non-Spatial Models*	0.046***			-0.003*	0.007***	0.287***
		(0.005)			(0.002)	(0.002)	(0.021)
	Spatial Error Model	-1.270***		0.474***	-0.495***	0.214***	0.404***
		(0.133)		(0.037)	(0.082)	(0.057)	(0.025)
	Spatial Lag Model	-0.122	-0.484***		-0.137***	0.026	0.146***
		(0.073)	(0.055)		(0.031)	(0.017)	(0.013)
	Spatial Lag Model + Spatial Error Model (ML)	-2.188***	0.876***	0.585***	-0.250***	0.076	0.145***
		(0.125)	(0.026)	(0.027)	(0.082)	(0.048)	(0.024)
	Spatial Model (GM)	-1.065***	0.357		-0.520***	0.177***	0.449***
		(0.143)			(0.085)	(0.058)	(0.025)

Source: Authors' elaboration based on software *R*.
Note: * p<0.10; ** p<0.05; *** p<0.01.
Standard deviations are in parenthesis under the coefficients.

Table 4. Price-Elasticity of Demand for Ethanol: Estimates for Brazil

Methods	*Models*	*Constant*	*Rho (ρ)*	*Lambda (λ)*	*Gasoline Price*	*Ethanol Price*	*GDP*
Pooled OLS	*Non-Spatial Models*	-0.007*** (0.003)			0.023*** (0.002)	-0.032*** (0.002)	0.334*** (0.013)
	Spatial Lag Model	0.531*** (0.020)	-1.855*** (0.202)		3.316*** (0.211)	-3.040*** (0.144)	0.584*** (0.034)
	Spatial Error Model	-4.236*** (0.249)		0.703*** (0.021)	3.920*** (0.264)	-3.705*** (0.162)	0.702*** (0.040)
Fixed Effect	*Non-Spatial Models*	0.000 (0.002)			0.007*** (0.002)	-0.019*** (0.002)	0.564*** (0.020)
	Spatial Lag Model		-0.035 (0.038)		2.828*** (0.280)	-1.730*** (0.157)	0.083*** (0.094)
	Spatial Error Model			-0.059 (0.040)	2.854*** (0.278)	-1.703*** (0.154)	0.086*** (0.093)
Random Effect	*Non-Spatial Models*	-0.001 (0.002)			0.008*** (0.002)	-0.019*** (0.002)	0.534*** (0.019)
	Spatial Error Model	-3.899*** (0.389)		0.647*** (0.027)	2.819*** (0.263)	-2.037*** (0.184)	0.774*** (0.081)
	Spatial Lag Model	-0.775*** (0.147)	-0.530*** (0.059)		0.377*** (0.069)	-0.362*** (0.038)	0.148*** (0.032)
	Spatial Lag Model + Spatial Error Model (ML)	-4.865*** (0.318)	0.972*** (0.572)	0.006*** (0.026)	1.268*** (0.186)	-1.113*** (0.107)	0.138** (0.054)
	Spatial Model (GM)	-3.243*** (0.408)	0.531		2.769*** (0.274)	-2.106*** (0.189)	0.961*** (0.080)

Source: Authors' elaboration based on software *R*.
Note: * p<0.10; ** p<0.05; *** p<0.01.
Standard deviations are in parenthesis under the coefficients.

Results of the estimation of price elasticity of demand for ethanol are shown in Table 4. The interpretation of spatial components in all the models is similar to those reported in Table 3. Regarding the results in Table 4, values of ϕ, of the spatial error and of spatial lag random effects models were significant at 1 and 5% statistical levels, respectively, but the value of ϕ in the spatial error+lag model was not significant at 10%. For those two models, specification of random effects is important, but not in the latter. The value of θ was equal to 0.848 for GM spatial random effects model, suggesting that the random effects model is appropriate.

Results of the auxiliary coefficients ϕ and θ of the spatial random effects models mentioned in the previous tables suggest that the random effects model is adequate to estimate the price elasticities of demand for gasoline and ethanol in Brazil. However, two tests were performed to provide more information to assist in the process of defining the most suitable model. These tests were developed and improved over the years (Baltagi *et al.*, 2003; Baltagi and Li, 2006; Baltagi *et al.*, 2007a; Baltagi *et al.*, 2007b; Baltagi *et al.*, 2007c; Baltagi and Liu, 2008). The first test is the Joint One-Sided Test to verify the joint significance of regional random effects and spatial correlation. The second test is the Conditional LM Test which tests for the presence of spatial correlation. Unlike the Unconditional Test, this test is performed on the results of the regressions.

These tests are important because of a relative lack of other tests that are more comprehensive. Thus, both tests are based on the random effects model. In the case of the Joint Test, such specification is tested, while in the Conditional Test this specification is assumed to be valid. Table 5 shows the results for these tests. The Joint Tests for gasoline demand and ethanol demand were both significant at 1%, which suggests that the specification of random effects model is adequate and that spatial correlation must be taken into account. Individual tests for the demands of both fuels were also significant at 1%, reinforcing previous results about the presence of spatial correlation. Therefore, the model that seems most appropriate is the random effects model, which incorporates both types of spatial correlation (spatial lag and spatial error), considering the ML estimation method.

Table 5. Baltagi, Song and Koh Tests for Regional Effects and Spatial Autocorrelation

Tests	*Description*	*Demand for Gasoline*	*Demand for Ethanol*
One-Sided Joint Test	*LM-H*	12931.990	4796.088
	p-value	1.00E-02	1.00E-02
Conditional LM Test	*LM-Lambda*	11.303	16.355
	p-value	2.2E-16	2.20E-16

Source: Authors' elaboration based on software *R*.

Considering the random effect model in Table 3, according to the econometric tests, parameters for the spatial lag + spatial error model estimated by ML show that gasoline demand is inelastic. Its price elasticity is -0.250. It has the expected sign and is close to -0.319 for the short-run and to -0.227 for the long-run elasticities estimated by Burnquist and Bacchi (2002). It also is close to -0.464 for the elasticities estimated by Alves and Bueno (2003). The five random effect models estimated presented the same sign, which shows robustness in this sense, although magnitudes were different. This may be due to the way in which spatial autocorrelation is treated by these models. In the case of the spatial lag + error spatial model, both the error component and lag are taken into account, better explaining the problem of spatial dependence, since both *lambda* and *rho* were significant at the 1% level.

Comparing the results of this model with other fixed effects and pooled OLS models, the sign is the same, only differing in the magnitude of coefficients. The cross-price elasticity regarding ethanol is 0.076. Although it presents the expected sign, it is not close to results presented in the relevant literature and it is not statistically significant for this model. However, its values are close to the five random effects models. Gasoline consumers are much less sensitive to ethanol prices than to those of gasoline. Finally, the income elasticity of 0.156 has the expected sign and shows that gasoline consumers are a lower sensitivity to income. This value is close to that presented in Alves and Bueno (2003), who found a value of 0.122 for the short and the long run and close to results presented in Roppa (2005), who found a value 0.163 for the long run.

The same random effects model in Table 4 shows that the demand for ethanol is elastic to prices. There is no reliable study in the literature to be used for comparison. The price elasticity of -1.113 has the expected sign, but its value is considerably high for the elasticity energy pattern. The same occurs to cross-price elasticity regarding the gasoline demand of 1.268, which has the expected sign, but a value that is also high for the elasticity energy pattern. This demonstrates that the consumption of ethanol has a very elastic demand in Brazil. Since the price elasticity of gasoline is lower in absolute values than that of ethanol, an increase in ethanol prices leads consumers to shift towards using gasoline more rapidly than they would in the opposite situation. Finally, income elasticity for an ethanol demand of 0.138 has the expected sign and also shows that ethanol consumers are less income sensitive. Elasticity of the ethanol demand curve regarding price is an evidence of the full competition in the fuel market, mainly in the ethanol market. The large amount of *flex-fuel* vehicles in the Brazilian automobile fleet results in a strong competition in this market.

5. Final Remarks

The objective of this paper was to estimate the price, cross-price and income elasticity of gasoline and ethanol in Brazil using spatial panel data models. The fuel market for light vehicles in this country is considerably competitive because of the fuel diversification and the *flex-fuel* technology. On the other hand, spatial features of the fuel supply and regional factors might determine heterogeneities in the behavior of regional consumers and in the existence of spatial autocorrelation patterns for fuel demand. Besides the estimation of elasticities, the role of individual and spatial heterogeneities in the per capita fuel consumption is also investigated. Spatial panel data models were used to estimate gasoline and ethanol demand equations for Brazil using a panel data set covering the 27 Brazilian states.

The spatial autocorrelation index (Moran's *I*) indicated that, in general, regions with a high per capita consumption of gasoline or ethanol were surrounded by regions this same consumption pattern. In addition, estimations showed that the coefficients of the pooled OLS spatial lag and error components, fixed-effect and random-effect models were all statistically significant. Spatial lag models suggested the existence of significant spatial spillover effects affecting the consumption of gasoline and ethanol. There also was an indication that it was important to consider serial correlation, leading to specification of the models by random effects. Results for the elasticities showed that gasoline demand was inelastic. Gasoline consumers were much less sensitive to ethanol prices than to gasoline prices, and also less income sensitive. Regarding ethanol, results showed that its demand is elastic to price. The price and cross-price elasticities were greater than 1 and considerably high for the energy elasticity pattern. This is a possible influence of *flex-fuel* vehicles, which increased competition in the fuel market in the last years. Income elasticity of the ethanol demand showed that ethanol consumers also were less income sensitive.

Despite the little time past from the application of spatial analysis to panel data, results showed that spatial panel data models provide an important contribution to estimate fuel demand equations. The control of spatial heterogeneity and individual heterogeneity might improve the accuracy of the estimations. Future research might be developed using a panel database that is less aggregated, depicting a higher number of spatial units, depending on availability of data.

References

Alves, D. C. O; Bueno, R. D. L. S. (2003). A. Short-run, Long-run and Cross Elasticities of Gasoline Demand in Brazil. *Energy Economics*, 25(2): 191-199.

Anselin, L. (1999a). *Spatial Data Analysis with SpaceStat and ArcView – Workbook*, Urbana-Champaign: University of Illinois, 3 ed.

Anselin, L. (1999b). *Spatial Econometrics*, Dallas: School of Social Science, University of Texas.

Anselin, L. (1992). *SpaceStat tutorial: a workbook for using SpaceStat in the analysis of spatial data*, Urbana-Champaign: University of Illinois.

Anselin, L. (2005). *Exploring Spatial Data with GeoDa*^{TM}*: A Workbook*. Center for Spatially Integrated Social Science, GeoDa Tutorial – Revised Version.

Anselin, L. (1988). *Spatial econometrics: Methods and Models*, Kluwer Academic: Boston.

Anselin, L., Bera, A. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. In: Ullah, A.; Giles, D. E. A. (Eds.), *Handbook of applied economic statistics*, Nova York: Marcel Dekker, pp. 237-289.

Anselin, L.; Le Gallo, J.; Jayet, H. (2008). Spatial Panel Econometrics. *Advanced Studies in Theoretical and Applied Econometrics*, 46, Part II, pp. 625-660.

Balestra, P.; Nerlove, M. (1996). Pooling cross-section and time-series data in the estimation of a dynamic model: The demand for natural gas. *Econometrica*, 34(3): 585-612.

Baltagi, B. H. (2001). *Econometric Analysis of Panel Data.* Chichester: John Wiley and Sons, Ltd.

Baltagi, B. H.; Bresson, G.; Pirotte, A. (2007a). Panel unit root tests and spatial dependence. *Journal of Applied Econometrics*, 22: 339-360.

Baltagi, B. H; Egger, P.; Pfafermayr, M. (2007b). Estimating models of complex fdi: are there third-country e effects? *Journal of Econometrics*, 140: 260-281.

Baltagi, B. H.; Griffin, J. M. (1983). *Gasoline demand in the OECD:* an application of pooling and testing procedures, *European Economic Review*, 22(2): 117-137.

Baltagi, B. H.; Li, D. (2006). Prediction in the Panel Data Model with Spatial Correlation: the Case of Liquor. *Spatial Economic Analysis*, 1(2): 179-185.

Baltagi, B. H; Liu, L. (2008). Testing for random effects and spatial lag dependence in panel data models. *Statistics and Probability Letters*, 78: 3304-3306.

Baltagi, B. H.; Song, S.; Jung, B.; Koh, W. (2007c). Testing for serial correlation, spatial autocorrelation and random effects using panel data. *Journal of Econometrics*, 140(1): 5-51.

Baltagi, B. H.; Song, S.; Koh, W. (2003). Testing panel data regression models with spatial error correlation. *Journal of Econometrics*, 117: 123-150.

Bentzen, J. (1994). An empirical analysis of gasoline demand in Denmark using cointegration techniques. *Energy Economics*, 16(2): 39-143.

Brazilian Institute of Geography and Statistic (IBGE). Data of price index and population. Available at: http://www.ibge.gov.br.

Brazilian National Institute of Oil and Gas (IPT). Data of Brazilian natural gás vehicles fleet. Available at: www.ibp.org.br.

Brazilian National Transportation Department (DENATRAN). Frota nacional de veículos automotores. Available at: www.denatran.gov.br.

Brazilian National Treasury. Tax Revenue by states. Available at: http://www.receita.fazenda.gov.br.

Brazilian Sugarcane Industry Association (UNICA). Ethanol Production. Available at: www.unica.com.br.

Burnquist H. L.; Bacchi, M. R. P. (2002). *A demanda por gasolina no Brasil: uma análise utilizando técnicas de co-integração.* CEPEA, Discussion paper. Available at: www.cepea.esalq.usp.br/pdf/DemandaGasolina.pdf.

Cliff, A. D.; Ord, J. K. (1981). *Spatial processes: models and applications*. Pion, London.

Dahl, C. and Sterner, T. (1991a). Analyzing gasoline demand elasticities*:* a survey. *Energy Economics*, 13(3): 203-210.

Dahl, C. and T. Sterner (1991b): ‘‘A survey of econometric gasoline demand elasticities’’, *International Journal of Energy Systems*, 11: 53-76.

Dahl, C. (1995). Demand for transportation fuels: a survey of demand elasticities and their components. *The Journal of Energy Literature*, 1(2): 3-27.

Elhorst, J. P. (2003). Specification and estimation of spatial panel data models. *International Regional Sciences Review*, 26(3): 244-268.

Elhorst, J. P. (2005). Unconditional maximum likelihood estimation of linear and log-linear dynamic models for spatial panels. *Geographical analysis*, 37(1): 85-106.

Elhorst, J. P.; Piras, G.; Arbia, G. (2010). Growth and convergence in a multi-regional model with space-time dynamics. *Geographical Analysis*, 42: 338–355.

Eltony, M. N.; Al-Mutairi, N. H. (1995). Demand for gasoline in Kuwait: an empirical analysis using co-integration techniques, *Energy Economics*, 17(3): 249–253.

Espey, M. (1996a). Explaining the variation in elasticity estimates of gasoline demand in the United States: a meta-analysis, *The Energy Journal*, 17(3): 49-60.

Espey, M. (1996b). Watching the fuel gauge: an international model of automobile fuel economy, *Energy Economics*, 18(1-2): 93-106.

Espey, M. (1998). Gasoline demand revisited: an international meta-analysis of elasticities, *Energy Economics*, 20(3): 273-95.

Grahan, D.; Glaister, S. (2002). The Demand for Automobile Fuel: a survey of elasticities, *Journal of Transport Economics and Policy*, 36: 1-26

Kapoor, M.; Kelejian, H.; Prucha, I. (2007). Panel data model with spatially correlated error components. *Journal of Econometrics*, 140(1): 97-130.

Lee, L.; Yu, J. (2010a). A spatial dynamic panel data model with both time and individual fixed effects. *Econometric Theory*, 26: 564-597.

Lee, L.; Yu, J. (2010b). Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*, 154: 165-185.

Le Gallo, J.; Ertur, C. (2003). Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe, 1980–1995. *Papers in Regional Science*, 82: 175-201.

LeSage, P. (1999). The Theory and Practice of Spatial Econometrics. University of Toledo, USA. Available at: www.spatial-econometrics.com/html/sbook.pdf.

National Agency of Oil and Biofuel (ANP). (2008). *Levantamento mensal de preços e vendas de combustíveis no Brasil*. Available at: http://www.anp.gov.br.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, 49(6): 1417-1426.

Pace, R. K.; Barry, R. (1997). Sparse spatial autoregressions. *Statistics and Probability Letters*, 33: 291-297.

Pinkse, J.; Slade, E. (1998). Contracting in space: an application of spatial statistics to discrete-choice models. *Journal of Econometrics*, 85: 125-154.

Polemis, M. (2006). Empirical assessment of the determinants of road energy demand in Greece, *Energy Economics*, 28(3): 2006.

Puller, S.; L. Greening (1999). Household adjustment to gasoline price change: an analysis using 9 years of US survey data, *Energy Economics*, 21(1): 37-52.

Ramanathan, R. (1999). Short and long-run elasticities of gasoline demand in India: an empirical analysis using co-integration techniques. *Energy Economics*, 21(4): 321-330.

Roppa, B. F. (2005). *Evolução do consumo de gasolina no Brasil e suas elasticidades: 1973 a 2003*. UFRJ: Rio de Janeiro.

Rouwendal, J. (1996). An economic analysis of fuel use per kilometer by private cars, *Journal of Transport Economics and Policy*, 30(1): 3-14.

Salvo, A.; Huse, C. (2010). Is Arbitrage Tying the Price of Ethanol to that of Gasoline? Evidence from the Uptake of Flexible-Fuel Technology, *Energy Journal*, 32(3): 119-148.

Yu, J.; de Jong, R.; Lee, L. (2008). Quasi maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and t are large. *Journal of Econometrics*, 146:118-134.