ECONOMIC DISCUSSION PAPERS

Efficiency Series Paper 4/2014

A latent class approach for estimating energy demands and efficiency in transport: An application to Latin America and the Caribbean

Manuel Llorca, José Baños, José Somoza, Pelayo Arbués

Departamento de Economía

Universidad de Oviedo

Available online at: http://economia.uniovi.es/investigacion/papers

A LATENT CLASS APPROACH FOR ESTIMATING ENERGY DEMANDS AND EFFICIENCY IN TRANSPORT: AN APPLICATION TO LATIN AMERICA AND THE CARIBBEAN

Manuel Llorca*
Oviedo Efficiency Group, Department of Economics, University of Oviedo

José Baños
Oviedo Efficiency Group, Department of Economics, University of Oviedo

José Somoza
Marine Research Centre, University of Havana

Pelayo Arbués
Oviedo Efficiency Group, Department of Economics, University of Oviedo

8 October 2014

Abstract

In this paper, we use a stochastic frontier analysis approach to estimate demand functions for energy in the transport sector. The estimation of frontier functions allows us to obtain energy efficiency measures that are a robust alternative to the energy intensity indicators that are commonly used for international comparisons. Due to the likely unobserved heterogeneity among countries, we propose the use of a latent class model that allows to test for the existence of groups of countries with clearly differentiated demands that are associated with distinct price and income elasticities. This study is the first to use a latent class stochastic frontier approach in the estimation of energy demand functions. The proposed procedure is applied to Latin America and the Caribbean, where the transport sector represents a large share of the total energy consumption.

These energy demand functions call for the inclusion of energy price in their estimation. As the transport of both goods and passengers implies the consumption of different types of energy, an index that aggregates various energy prices is thus required for the analysis. However, international agencies do not provide specific indicators of aggregate energy prices in transport for the majority of the countries analysed. For this reason, in this paper we propose the construction of a transitive multilateral index which, in contrast to those frequently presented by the aforementioned agencies, facilitates international comparisons over time.

Keywords: energy demand in transport; efficiency; latent class stochastic frontier model; transitive multilateral price index; Latin America and the Caribbean.

---

* Corresponding author: Faculty of Economics and Business, University of Oviedo, Av. del Cristo, s/n, 33006, Oviedo, Spain. Tel. +34985104885. E-mail: llorcamanuel@uniovi.es

Acknowledgements: This research is partially funded by the Government of the Principality of Asturias through its predoctoral grants program.

1. Introduction

Since the 1970s oil crisis, the measurement and control of energy efficiency has become an essential goal of the economic and energy policies of a large majority of countries, especially in those that import energy. This interest subsequently arose in the late 1980s as a result of the growing awareness of global warming. A key issue in the strategy of the countries that aim to reduce their energy consumption and mitigate their greenhouse gas emissions is the adoption of measures that improve the efficiency of energy use in all economic sectors and especially in those that are energy intensive, such as transport.

Figure 1 shows that transport is the sector with the highest energy consumption in Latin America and the Caribbean. In recent decades, this sector used, on average, 43% of the total energy consumption, followed by manufacturing at 37%, household consumption at 14% and the service sector at 6%. The Economic Commission for Latin America and the Caribbean (ECLAC, 2010) indicates that the transport of passengers and goods will increase in the future. Combined with the dissociated manner in which public policy on infrastructure and transport has been conducted, this will result in an increase in the future use of energy, which implies a significant amount of oil derivatives consumption in the near future. Nevertheless, little published information on the transport sector in Latin America and the Caribbean is available. It is thus necessary to conduct studies focused on the energy consumption of this sector that can help to mitigate the environmental sustainability issues that are mentioned in the "Millennium development goals" proposed by ECLAC (2005).

[Insert Figure 1 here]

Per capita energy consumption in Latin America and the Caribbean is currently low in comparison with other world regions. However, since the 1990s, it has experienced significant growth, as shown in Figure 2. This low per capita consumption does not necessarily indicate high efficiency in the use of energy, as a significant part of the population of these countries lack the funds to have access to a private car. In this context, the rapid development of the region in the medium term might lead to unsustainable increases in the energy consumption of the transport sector and to the associated emissions of greenhouse gases. For example, between 1990 and 2007, the vehicle fleet that was used in Brazil, Mexico, Chile and Colombia increased by 53 million vehicles (the amount tripled), with 40% of this increase concentrated between 2003 and 2007. Therefore, it is crucial to elaborate orderly development strategies that favour public transport[1] and promote energy efficiency.

[Insert Figure 2 here]

Figure 3 shows the change in energy price in the transport sector between 1990 and 2010. This price is a transitive multilateral index that the current authors have elaborated. It adds the weighted prices of the different types of energy that are consumed in transport (See Appendix). The scenario of low energy prices in the 1990s contrasts with the inflationary process that was experienced in the first decade of the 21st century. This process also led many Latin American countries, especially those that were net importers of energy, to adopt programs to improve energy efficiency. These measures aim, on the one hand, to modernize public transport to incentivize its use,

[1] The lack of public services can lead to inefficient individual consumption decisions. The deficit (in quantity, quality or both) in the public transport services, for example, incentivizes private transport, which may generate high costs for the user and cause contamination and congestion in the cities.

renovate the vehicle fleets, introduce biofuels as alternatives to oil, promote the use of hybrid and electric vehicles, and promote the use of trains and subways in certain activities. On the other hand, the infrastructure network should be improved in tandem with logistical solutions to the provision of services, such as the adoption of intelligent measures that optimize transport routes and favour intermodality (ECLAC, 2010).

[Insert Figure 3 here]

An additional issue that should be addressed is the review of subsidy policies on transport and on products derived from oil, with the aim of transmitting adequate signals to the economy and achieving improvements on energy efficiency. This goal has frequently clashed with the pressure that has been exerted by consumers from various countries who rally against increases in energy prices. In that sense, Latin American countries should introduce a fiscal, incentive and environmental regulations system similar to the ones that exist in other parts of the world, such as the European Union and the United States (Barros and Prieto-Rodríguez, 2008; Chavez-Baeza and Sheinbaum-Pardo, 2014).

To help achieve the goal of reducing energy consumption, various quantitative indicators that are related to the energy efficiency of each country have been developed and have been used in international comparisons. There is no single definition totally accepted for the concept of energy efficiency, both in terms of the economy as a whole or specifically the transport sector. However, Ang (2006) and Stead (2001) indicate that the most common practice has been to link this idea with some thermodynamic, physical-based and monetary-based indicators that relate energy consumption to measurements of the economic activity or energy services derived from this consumption. The most commonly used indicator is the ratio of energy consumed to Gross Domestic Product (GDP). This measure of energy intensity has the advantage of simplicity in its calculation and easy interpretation, thus leading to its continued use in international statistics of the International Energy Agency (IEA), the World Bank and the World Energy Council. Decreasing levels of this indicator represent, on average, a reduction in the energy that is required to generate a unit of national production. Therefore, energy intensity is simply the inverse of the energy productivity indicator. Nonetheless, the value of energy intensity can vary significantly over time due to the changes in the structure of GDP, which are difficult to assimilate into the concept of energy efficiency.[2] Furthermore, these types of measures are not "relative", i.e. do not allow cross-country comparisons for countries with better practices or the calculation of potential energy savings.

The main goal of this paper is to adapt the methodological proposal of Filippini and Hunt (2011, 2012), for energy consumption in the transport sector of Latin America and the Caribbean. This adaptation is performed to obtain a relative measure of energy efficiency that overcomes the weaknesses of other indicators and can serve for international comparisons that are consistent throughout time. The measures of energy efficiency that are obtained are bounded and allow for the determination of potential energy savings given the characteristics of a country. Furthermore, the current study estimates various functions of frontier demand using a latent class approach. This type

[2] The IEA (2014) recognises that the use of energy intensity as a proxy for energy efficiency can generate untrustworthy results. Despite significant interest in the measurement of energy efficiency, its calculation for the transport sector is a difficult task. This organization proposes indices of energy intensity for the sector that are calculated using various disaggregated indicators that are obtained from large quantities of information. Due to this requirement, it is impossible to calculate this measure for all Latin American and Caribbean countries.

of approach takes into account the potential existing heterogeneity among the countries analysed, obtaining different demands that are associated with specific price and income elasticities for different country groups. To the best of our knowledge, this study is the first to apply this type of methodology for both the transport sector and the Latin American countries.[3]

This paper is organized as follows. In Section 2, we define the general demand for energy in the transport sector by providing a brief review of the existing literature. Additionally, we propose the use of a Stochastic Frontier Analysis (SFA) approach and the application of a latent class model. In Section 3, we present the database and the econometric specification of our models. The results of the estimations are presented in Section 4 and finally, Section 5 ends the paper with a summary and the presentation of conclusions.

2. Energy demand of the transport sector

Transport demand is derivative in nature, as the goal of moving goods and people is not to perform the journey but to reach a certain destination. In other words, demand is derived from the mobility of passengers and goods. This mobility, in turn, leads to energy or fuel demand, which is necessary for transport.

The previous research in the literature on the modelling of energy consumption for transport can be clustered into works that apply econometric techniques, those that use artificial intelligence approximations, those that use multi-criteria analysis and those that employ simulation methods (see Limanond *et al.*, 2011 or Suganthi and Samuel, 2012 for a review). The first group includes multiple linear regression models (Limanond *et al.*, 2011), partial least square regressions (Zhang *et al.*, 2009) and the analysis of time series and cointegration (Samimi, 2003; Galindo, 2005; Sa'ad, 2010; Hao, 2011). The second group includes studies of artificial neural networks (Dreher *et al.*, 1999; Murat and Ceylan, 2006; Limanond *et al.*, 2011) and harmony search algorithms (Haldenbilen and Ceylan, 2005; Ceylan *et al.*, 2008). Some studies have even combined the analysis of time series and fuzzy logic (Al-Ghandoor *et al.*, 2012). In the prediction of energy consumption for vehicles, the use of multi-criteria analysis should be noted, such as in the works of Lu *et al.* (2008) and Lu *et al.* (2009). Lastly, the most prominently used simulation model has been the Long-range Energy Alternatives Planning System (LEAP), which allows planning alternative scenarios for energy demand in the transport sector. The works that utilized this method include Bauer *et al.* (2003), Manzini (2006), Pradhan *et al.* (2006) and Islas *et al.* (2007).

Therefore, there is an extensive body of literature on the economics of transport that estimates various functions of energy consumption or the respective functions of fuel use for different types of vehicles. These studies have typically aimed at predictive purposes. The current study belongs to the line of econometric approximations of energy demand from the transport sector that calculates the price and income elasticities that are related to energy consumption (see, for example, Dahl, 1995). In their literature review, Graham and Glaister (2002) observe that, as a general rule, price elasticities that

[3] The scarcity of empirical analyses in this context has been conditioned by the availability of statistics. In fact, in many Latin American countries, there is no formal link between institutions that are in charge of providing information on energy and transport. Consequently, in this paper, all variables that are relative to energy consumption are based on the author's own work on the data provided by the Latin American Energy Organization (OLADE in Spanish).

are obtained in the short term are commonly between -0.2 and -0.3 and that those obtained in the long term are between -0.6 and -0.8. For the case of the income elasticities, they find that are often greater than one (between 1.1 and 1.3) in the long term and between 0.35 and 0.55 in the short term. The papers that are included in their review generally analyse Organization for Economic Cooperation and Development (OECD) countries. Wohlgemuth (1997) presents elasticities for several countries that are not OECD members. In terms of Latin America and the Caribbean, the elasticities for Mexico[4] and Brazil are presented. In the long term, the income elasticities for Mexico are between 0.99 and 1.72 and the price elasticities are between -0.04 and -0.21. For the case of Brazil, the income elasticities are between 0.88 and 1.10 and the price elasticities are between -0.10 and -0.26.

In general, in the traditional transport literature, energy demand is understood as a standard demand function. As previously mentioned, in the proposal that is presented below, a stochastic frontier function, which is similar to the production/cost functions that are commonly estimated in efficiency and productivity studies, is considered.

2.1. A stochastic frontier approach for energy demand in transport

A generic function of energy demand, which positively depends on income and inversely depends on prices, can be presented in the logarithmic form as follows:

$$\ln Q = \ln f(P, Y, X, \beta) + \varepsilon \qquad (1)$$

where Q represents the quantity of the demanded energy, P is the price of energy, Y represents income, X refers to other control variables, β are the parameters that are associated with the variables that are included in the model and can be directly interpreted as elasticities, and ε is the random error, which is commonly assumed to follow a normal distribution with a mean of zero and constant variance, σ_ε^2.

This assumption for the stochastic part of the function indicates that the researcher assumes that any deviation in energy demand that is predicted by the deterministic part of the model is a result of random shocks such as measurement errors or uncertainty. Therefore, this model can be estimated using the common estimator of Ordinary Least Squares (OLS), which allows for consistent and unbiased estimates of the model parameters under certain assumptions.

Although this approach has traditionally been used in empirical work, it does not provide direct information on one of the main issues of interest in the field of energy consumption in recent decades, i.e., energy efficiency. As stated in the previous section, there has been debate about the definition and measurement of this concept. In essence, this concept attempts to capture the relation between energy consumption and the production or service that is derived from this consumption. It should be measured in such a way that an improvement in the indicator implies a lower use of energy to produce a certain amount of output in a given economy.

However, in contrast to the research in the energy economics literature, the production economics field has developed various approaches that allow for the inclusion of efficiency in the activities of companies (or countries) within the random part of the model, without the need to add new variables or rely on other indicators.

[4] Although in the paper of Wohlgemuth (1997) Mexico is included in the group of countries that are not members of the OECD, this country was already a member since May 18, 1994.

Based on the efficiency and productivity literature, Filippini and Hunt (2011, 2012) suggest the use of a parametric approach of stochastic frontiers to estimate aggregate energy demand functions that are derived from a cost function in the provision of energy services. In this cost function, energy is an input. Thus, following Shephard's lemma and deriving the function based on the price of energy, the demand function of this input can be obtained. The main goal of those authors is to obtain measures of energy efficiency that can be used as alternatives to the typical indicators of energy intensity. These efficiency measures are based on the comparison of the energy consumption of the countries with respect to the minimal energy consumption predicted by the frontier, which takes into account the optimizing behaviour of companies and individuals.

The basic model that is estimated by those authors is the standard SFA model that was initially proposed by Aigner, Lovell and Schmidt (1977) (hereinafter ALS), but they also estimate other models developed in the efficiency and productivity literature, such as the True Random Effects model (TRE) presented by Greene (2004, 2005a, 2005b) or the formulation of Mundlak (1978) that was proposed for an estimator of random effects by Farsi *et al.* (2005). The standard ALS model can be presented for the case of energy demand as follows:

$$\ln Q = \ln f(P, Y, X, \beta) + v + u \tag{2}$$

where the random term can be decomposed in v, which is a normal distribution that is analogous to that represented by ε in equation (1), and u, which is an asymmetric error that follows a half-normal positive distribution to capture the inefficiency of energy demand. In the SFA literature, it is typically assumed that u is a negative half-normal (or truncated normal) if the function that is estimated is a production function with a maximum achievable production and positive if the estimated function is a cost function with an achievable minimum cost. In the case of a frontier demand, such as the proposed by Filippini and Hunt (2011, 2012), efficient energy demand represents a minimum feasible consumption. Thus, the approach that is used is the same as that for a cost function.

Based on the conditional mean of the inefficiency term proposed by Jondrow *et al.* (1982), the efficiency level for each observation can be obtained by applying the following expression:

$$EF_{it} = \frac{Q_{it}^{*}}{Q_{it}} = \exp\left(-\hat{u}_{it}\right) \tag{3}$$

where Q_{it}^{*} represents the aggregate energy demand of the country i in period t on the frontier, i.e., the minimum level of energy necessary for this economy to produce its output level; Q_{it} is the aggregate energy demand that is actually observed in this country; and EF_{it}, is thus a measure of efficiency that is bounded between zero and one. The difference between 1 and this measure of inefficiency shows the amount of energy consumption that could be reduced in this country (expressed as a decimal fraction) while maintaining the same level of transport services. Therefore, these are relative measures that, in contrast to energy intensity indicators, allow for direct comparisons between countries throughout time.

To explain the concept of stochastic frontier, in Figure 4, we compare various approaches that could be used in the econometric estimation of energy demand functions. The blue line shows the energy demand function that is proposed in Equation

(1) as estimated using OLS. With this approach, we obtain a function with a negative slope in relation to the prices that pass through the mean of the observed values. A basic frontier model that would allow the identification of countries' efficiency would simply assign the whole estimated error (i.e., $\hat{\varepsilon}$) that was obtained from applying OLS to the inefficiency. This simple approach does not allow the separation of inefficiency from noise because, by definition, a deviation from the minimum possible consumption that can be achieved is attributed to inefficiency. This type of frontier is typically known as deterministic frontier and can be obtained by moving the intercept of the OLS estimation until all observations are to the right of the estimated frontier. This form of frontier attainment is known as Corrected Ordinary Least Squares (COLS). In other words, it allows for the attainment of a function that envelops all observations. In the current case, it is represented by the blue dashed line. Although Filippini and Hunt (2011, 2012) do not represent it graphically, the demand that is estimated when an SFA approach is used, is a function such as that represented by the green line. The use of this type of methodology allows for certain observations to be to the left of the estimated frontier due to the existence of negative random shocks, although the majority of observations are to the right of the frontier due to the inefficiency effect.

[Insert Figure 4 here]

Figure 5 represents the type of frontier that is estimated when using the SFA approach to obtain energy demand and how the Overall Random Error (ORE), i.e., the stochastic part of the model (v+u), can be decomposed into inefficiency and noise[5] in the various possible cases. As shown for observation 1, an observation lies only at the frontier that we have presented graphically when the inefficiency term compensates the negative value of the noise term (or both are equal to zero). By estimating a stochastic frontier demand, it is assumed that the majority of observations will be located to the right of the frontier. This can be due to an effect either of inefficiency or noise (if it is positive and inefficiency is equal to zero), as in observation 2, or to both of these effects together, as is the case in observation 3 (in which both are positive) and 4 (in which only one part of the inefficiency is compensated by the negative value of the noise term). Nevertheless, as this is not a deterministic frontier, some observations can lie to the left of the estimated frontier, indicating that these countries use less energy than is predicted by the frontier for a specific price. Observation 5 is to the left of the estimated frontier because there is no inefficiency and the error term is negative. In observation 6, even with the existence of inefficiency, the negative noise term exceeds the value of u and thus, this observation is "super-efficient".

[Insert Figure 5 here]

2.2. Treating unobserved heterogeneity with a latent class model

Based on the influential work by ALS, a broad body of literature has been developed to attempt to precisely measure the efficiency of the studied individuals (firms, countries, etc.) with various methodological proposals that allow for solving specific problems that affect the obtained results. One of the main weaknesses of the

[5] The random component v includes events which cannot be controlled by transport companies or the individuals who use the private vehicle, such as those caused by the weather or natural disasters. Thus, by considering for Latin American countries the average amount of days per year when their transport infrastructure (roads, bridges, etc.) is cut off due to these causes, a deviation greater than (lower than) this value in a given year would produce a positive (negative) shock.

basic model that is proposed in equation (2) is that despite the fact that its specification allows to control for random noise, the presence of unobserved heterogeneity between the studied individuals can bias the efficiency measures (see Greene, 2005a, 2005b).

This heterogeneity is typically considered an unobserved determining factor of the estimated production or cost frontier, and inefficiency is interpreted as the distance to the frontier once heterogeneity has been taken into account. Multiple empirical strategies, each with specific advantages and drawbacks, have been developed to solve this problem. A first approach that can be applied and that is commonly used as an analogy to the traditional econometric literature is the use of a specification that includes individual effects (fixed or random), as is the case for the True Fixed Effects (TFE) and TRE models proposed by Greene (2004, 2005a, 2005b). These models include a series of country-specific intercepts that are simultaneously estimated with remaining parameters of the model and allow the distinction between unobserved heterogeneity (which does not change over time) and inefficiency. In this approach, unobserved heterogeneity additionally enters the model as an individual-specific intercept and, therefore, is a neutral or parallel movement of the function that maintains the remaining common parameters for all individuals. In the case of energy demand, as estimated in the current paper, this implies that specific characteristics of this demand, such as their price and income elasticities, are the same for all countries analysed. This assumption is difficult to justify for such a heterogeneous region as Latin America and the Caribbean. If there are different groups of countries in the sample with different demand characteristics, i.e., different parameters that are associated with the variables, we should estimate a model that allows us to take this feature into account.

An alternative approach to control for unobserved heterogeneity that seems to be adequate for the current context is the Latent Class Stochastic Frontier Model (LCSFM), such as that proposed by Orea and Kumbhakar (2004) and Greene (2004, 2005b). This model allows for estimation of different parameters for countries that belong to distinct groups and share similar characteristics. The characteristics of the countries in each group differ and thus, given that the countries that belong to the same class share the same set of parameters, this approach controls for the existing heterogeneity between the groups. In other words, the latent class procedure allows us to control for heterogeneity in the slopes (the coefficients of the estimated variables), which is unobserved and associated with country groups. The estimation of a model of this type implies the existence of *J* groups of countries, which demonstrate differences between themselves in terms of their behaviour function:

$$\ln Q_{it} = \ln f\left(P_{it}, Y_{it}, X_{it}, \beta_j\right) + v_{it}\big|_j + u_{it}\big|_j \tag{4}$$

where the subindex $j = 1,\ldots,J$ refers to class, β_j is the vector for the parameters that are estimated for class j, and the random term, as in prior models, is composed of $v_{it}|_j \sim N(0,\sigma_v^2|_j)$ and $u_{it}|_j \sim N^+(0,\sigma_u^2|_j)$, which are also specific for each class. The estimation of this model implies the maximization of the overall likelihood function from Equation (4), which is the sum of the likelihood functions at each point of the sample weighted by the probability of belonging to each class. This, in turn, is parameterized as a multinomial logit model. Additional variables can be included in the probabilities of class membership. If such variables are not included, the model uses the goodness of fit of each class to identify the distinct groups.[6]

[6] The estimation procedure of this type of model can be found in detail in Orea and Kumbhakar (2004) and Greene (2005b).

A necessary condition for the identification of the parameters of class membership probabilities in these models is that the sample has been generated by various behavioural functions or different error terms. In other words, J, the number of classes, is considered to be given. If J is greater than the "actual" number of classes and, thus, we attempt to estimate a model with "too many" classes, the model will be overspecified and the model parameters will not be able to be estimated. Thus, the researcher should select the number of classes J a priori. Various statistical tests can be used to choose the appropriate number of classes. The most commonly used tests are the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), and some of their variants. For these tests, the selected model is that with the lowest criterion value. Both of these criteria (and those derived from them) seek a balance between the lack of fit (by estimating a model with a small number of classes) and overfitting (by estimating a model with excessive classes). With that aim, apart from incorporating the value of the likelihood function, these criteria penalize (with different weights) the increase in the number of parameters that are estimated in each model. The BIC most severely penalizes the overfitting and tends to favour more parsimonious models than does the AIC. In this paper, we use various criteria to select the preferred model.

After the model estimation, the posterior probabilities can be obtained to assign each country to a specific class and calculate the efficiency measures. One strategy to assign countries is assuming that the country belongs to a class to which it may belong with the highest probability. Therefore, only one of the demands is taken as the reference frontier to obtain the (in)efficiency measure for each country.[7] An alternative method, as Greene (2005b) proposes, is to take all classes into account when calculating country efficiency, i.e., adding the specific efficiencies of belonging to each of the classes weighted by the probability of belonging to them. However, in this article, we use the first approach with the understanding that groups of countries actually have different demands.

3. Data and econometric specification

This section presents the data and the econometric specification of the models to be estimated that were presented above. Incomplete panel data are used, for the 1990-2010 period, from the following 24 countries in Latin America and the Caribbean: Argentina, Barbados, Bolivia, Brazil, Chile, Colombia, Costa Rica, Ecuador, El Salvador, Granada, Guatemala, Guyana, Honduras, Jamaica, Mexico, Nicaragua, Panama, Paraguay, Peru, Dominican Republic, Suriname, Trinidad and Tobago, Uruguay and Venezuela.[8] The econometric specification of the basic model (ALS) is the following:

[7] In this paper, we estimate the models assuming a panel data structure, i.e. the probabilities of belonging to each class are constant over time for each country. Therefore, each country is assigned to a single group throughout the sample period.

[8] The sample is composed of a total of 503 observations. The observation for Barbados in 2010 is not included because it is unavailable. Of the 27 country members of OLADE, Belize and Haiti are not included due to lack of information. Furthermore, Cuba is not included in the sample, as the inclusion of this country in the analysis does not allow for the convergence of estimates in some models because the estimated function does not fulfil the convexity property and, in other models, the obtained values for efficiency are practically zero. Due to these results, the observations for this country are considered to be outliers and, thus, we exclude them from the sample.

$$\ln Q_{it} = \alpha + \beta_Y \ln Y_{it} + \beta_{POP} \ln POP_{it} + \beta_P \ln P_{it} + \beta_{ST} ST_{it} + \beta_{DEN} \ln DEN_{it} + \beta_t t + \frac{1}{2}\beta_{tt} t^2 + v_{it} + u_{it} \quad (5)$$

where Q, Y, P, v, u and β are defined as in the prior equations. Analogously to Filippini and Hunt (2011, 2012), we include other explanatory control variables such as *POP*, which represents the population; *ST*, which is the share of the transport sector in the economy; *DEN*, population density; and *t*, the time trend, which is also introduced squared.[9]

Table 1 shows the descriptive statistics of these variables. It should be mentioned that the dependent variable, *Q*, represents the final energy consumption of the transport sector, expressed in thousands of toe. It is obtained by adding the total of the energy consumption in internal transport[10] for each country for both passengers and goods. The types of energy that are included in this aggregate are natural gas, liquid gas, electricity, gasoline (which includes biofuel), kerosene (jet fuel), diesel oil and fuel oil. *Y*, is the GDP of each country and is measured in millions of 2005 US dollars at Purchasing Power Parity (PPP). In international analyses, the use of this exchange rate is indispensable for adequately comparing the GDP across countries. *POP* is the mean population for each country, as measured in thousands of inhabitants. *P* is an energy price index in the transport sector, calculated as the weighted sum of mean prices of the types of energy used in the sector. Because OLADE and other energy international agencies do not provide any price index for the total of the countries of Latin America and the Caribbean, we have calculated a transitive multilateral price index that allows for consistent comparisons between countries throughout the sample period (see Appendix). *ST* is the ratio of Gross Value Added (GVA) in transport and the total GVA for each economy, and it is expressed in percentages. Lastly, *DEN* reflects the ratio between the population in thousands of inhabitants and the area of each country in km^2. This variable and the income per capita (*Y/POP*) are also included in the LCSFM model within the class membership probabilities to help with the segmentation of the sample.[11] Concerning the data sources, the variables *Q*, *P* and *POP* are derived from the Energy-Economic Information System of the OLADE. The variables *ST* and *DEN* are obtained from ECLAC. The variable *Y* is obtained from the data in the Penn World Table (PWT 7.1) presented by Heston *et al.* (2012).

[Insert Table 1 here]

If we pay special attention to the quantity and price of the consumed energy (i.e., the most relevant variables in a demand analysis apart from income), significant differences between countries can be observed. Figure 6 shows that energy consumption in transport for Latin America and the Caribbean evidenced a significant dynamism during the period analysed, with an average annual growth of 4.1%, which is more than double that of the growth in the UE-25 (1.3%) for the same period. Nevertheless, the

[9] However, in this paper, we do not include meteorological variables because we analyse energy demand in the transport sector and such variables do not play a relevant role as in the modelling of total energy demand or the residential sector of a country. However, possible persistent meteorological differences would be controlled for in the latent class model, which precisely allows the treatment of unobserved heterogeneity.

[10] Internal transport includes domestic aviation, domestic shipping, roads and railways and excludes international maritime and air transport.

[11] The lack of homogenous information or a sufficient timeframe on the transport infrastructure, stock of vehicles, distances travelled or goods and passenger traffic indicators, impedes the inclusion of these types of variables in the estimated demands.

growth rates were quite different among countries, with Jamaica and Suriname displaying the highest growth and Argentina and Colombia presenting the lowest growth for the period of analysis.

[Insert Figure 6 here]

In Figure 7, we represent the time evolution of the price index for the group of countries that evidenced the greatest differences in 2010. It should be noted that during the years analysed, Venezuela persistently maintained the lowest prices. Furthermore, it is noteworthy to mention the low cost of energy in Ecuador and Mexico. By contrast, the highest prices are found in Colombia, Bolivia, Brazil and Argentina. Furthermore, we observe that the price index that is used in this paper does not require that its own value be equal for all countries in a base year, which is required when standard indices such as Laspeyres or Paasche are applied. As discussed in the Appendix, this facilitates a better fit of the estimated energy demand functions.

[Insert Figure 7 here]

4. Estimates and results

Table 2 shows the results of the basic ALS model estimation. As previously mentioned, the model assumes the existence of a single demand and, therefore, does not allow for different elasticities for the various countries in the sample. All of the variables that are included in the models are statistically significant at 99% (except the time trend squared) and have the expected signs. The values of the income and price elasticities are 0.81 and -0.23, respectively. These elasticities are found within the value ranges that are obtained in the energy demand in transport papers, as discussed in Section 2. The coefficient of the population variable has a positive sign, which indicates (as expected) that a population increase leads to, *ceteris paribus*, an increase in the energy demand. A similar interpretation can be made for the share of the transport sector in the economy, which can be understood as a proxy for the degree of transport development. It can be expected that a more developed sector results in greater welfare for society, which is achieved through greater energy consumption. However, density presents a negative sign, indicating (as expected a priori) that the countries that are more densely populated have, *ceteris paribus*, lower transport energy demand due to the smaller average distances that companies and individuals travel. After controlling for the remaining variables in the estimation, the positive sign of the time trend shows that energy consumption increased throughout the sample period (as shown in Figure 2), which may indicate technical regress in the sector.[12] The mean value of efficiency is 87.4%. Nevertheless, great variability is found among the observations, with minimum and maximum values of 66.2% and 94.7% respectively.[13]

---

[12] This model has alternatively been estimated by including a set of time dummies that capture the non-linear evolution of energy consumption over time. Nevertheless, we prefer the inclusion of a time trend and its square, as it allows the estimation of a latent class model without renouncing the inclusion of the time effect in the model.

[13] A reviewer's suggestion that the inefficiency in our model might include a behaviour that would be the consequence of low energy prices in certain countries led us to estimate a heteroscedastic model of the type proposed by Reifschneider and Stevenson (1991), Caudill and Ford (1993) and Caudill, Ford and Gropper (1995). The coefficients that are estimated with such a model for the variables in the frontier are practically identical to those obtained in the ALS model, and the price is not statistically significant in the inefficiency term. The values of the efficiencies that are obtained in this heteroscedastic model are similar to those obtained in the ALS model, with a 96% correlation between the two measures.

[Insert Table 2 here]

Table 3 shows the results of the LCSFM models for two and three classes[14], which include separating variables in the probabilities of class membership. If we analyse the prior probabilities of the two-class model, the separating variables (income per capita and density) are not statistically significant. Thus, this model is equivalent to a model that does not include separating variables. However, these variables are significant in the three-class model. These variables' signs and values indicate that countries with higher income per capita and lower population density tend to be assigned to class 1 and, to a lesser degree, to class 2.

[Insert Table 3 here]

Figure 8 shows the different information criteria that are used as selection tests to choose the preferred model: the traditional AIC and BIC and some of their variants, the modified AIC criterion (AIC3), the corrected AIC (AICc), the AICu and the consistent AIC (CAIC), which can be considered a variation of the AIC and the BIC.[15] As mentioned in the previous section, all of these criteria are based on the maximum value of the likelihood function, which is obtained by estimating each model. These criteria only differ in that they penalize the increased number of parameters that are estimated for each model with different weights. The model with the best fit is that with the lowest criteria value. All of the presented criteria show a clear improvement in the fitness of the estimates when unobserved heterogeneity is addressed in the model through a latent class approach. Although a great improvement can be observed when moving from the ALS model to the LCSFM model with two classes (all criteria present a lower value), this heterogeneity is captured to an even greater extent by a three-class model. For the case of LCSFM models, the criteria values are also shown when these models are estimated without the inclusion of separating variables, although the estimated parameters are not presented in this paper. In the model with two classes, no improvement is observed when separating variables are included. By contrast, in the model with three classes, these variables have a relevant influence. This three-class model is the one that fits best to the characteristics of our data and, thus, we consider it to be the preferred choice.

[Insert Figure 8 here]

Table 3 shows that the majority of variables in these models are significant and have the expected signs, as in the ALS model. The preferred three-class model shows large differences in the coefficients between the classes for most of the variables. For example, the population variable coefficient varies between 0.247 and 0.742, the share of the transport sector in the economy is only significant for class 1, population density positively affects[16] energy consumption in group 2 and negatively affects it in group 1, and although all of the classes demonstrate a positive time trend, this growth in energy consumption is increasing in class 2 and decreasing in class 3 according to the sign of the square term in both cases.

---

[14] The model estimates with higher numbers of classes do not converge.

[15] Additional details on these criteria can be found in Fonseca and Cardoso (2007).

[16] As previously mentioned, a negative coefficient for *DEN* is expected because this variable mainly captures the effect of greater energy consumption as the territory of a country increases given its population. However, the coefficient of this variable is positive in class 2 of the latent class model. This result does not invalidate our intuition on this variable, as this ratio simply includes population divided by area and does not incorporate the degree of urbanization or whether the population is distributed homogeneously in the territory, a circumstance that may condition this result for this class.

The most relevant variables in demand analysis are income and price. As previously mentioned, the latent class model allows us to identify three classes with elasticities that clearly differ. The most inelastic demand for income can be found in class 2 (0.179), followed by class 1 (0.566) and finally, the most elastic class is 3 (0.649). The differences in price elasticities of the demand are also evident if we represent the data from the sample without performing any type of estimation and only use the partition of the sample that is generated by the preferred three-class model, as can be shown in Figure 9.[17] This figure shows that the demand of class 1 has the steepest slope and corresponds to the group with the lowest elasticity (-0.161) in the estimates. This group includes Argentina, Brazil, Chile, Ecuador, Guyana, Mexico, Paraguay, Suriname, Trinidad and Tobago and Venezuela. The demand of class 2 corresponds to the group of intermediate elasticity (-0.288) and is composed of Barbados, Bolivia, Colombia, Costa Rica, Jamaica and Panama. Finally, class 3, with the flattest slope in the graph, is the most elastic in the estimates (-0.407) and includes El Salvador, Granada, Guatemala, Honduras, Nicaragua, Peru, the Dominican Republic and Uruguay. This figure also represents the single demand that would be obtained if we did not take into account the heterogeneity between countries, thus obtaining a biased demand with an intermediate slope between class 2 and class 1, which would correspond to the price elasticity value obtained from the ALS model (-0.229).

[Insert Figure 9 here]

The mean efficiencies that are obtained in each class are around 95%, and the minimum value is consistently greater than 80%. These results indicate that the groups are more homogenous than when one single demand is estimated. These results reflect that more efficient countries can reduce their energy consumption up to 5% and that the less efficient countries have a margin of up to 20%.[18] The estimation of this latent class model allows us to identify the most efficient countries in each class (on average for the period considered). The remaining countries in each group, given their similar characteristics, should attempt to imitate these most efficient countries' energy policies. The two countries with the greatest energy efficiencies are Brazil and Mexico in class 1, Barbados and Colombia in class 2, and El Salvador and Guatemala in class 3.[19]

[17] In this figure, we present price to energy divided by income. This consideration allows us to "relativize" the weight of income and isolate, to a certain degree, the price effect on demand, which is the current aim. On the other hand, logarithms in the units of both axes are calculated to reduce the measurement scale and facilitate the representation of the demand.

[18] These potential savings are however obtained without taking into account possible "rebound effects". This phenomenon basically captures the idea that part of the savings from increases in the efficiency level in the use of energy can be offset by increases in the demand for energy services derived from the marginal cost reduction of those energy services. This concept has recently received great attention in energy studies and especially in transportation (see for instance, Greene *et al.*, 1999; Small and Van Dender, 2007; or Hymel *et al.*, 2010).

[19] The reference countries for each of the demands seem to correspond to the countries that, according to ECLAC (2010), have adopted distinctive measures for the improvement of public transport in their cities. In this report, it is highlighted the Rapid Transit Bus (RTB) system implementation in Curitiba (Brazil). This system was started in 1972 as part of a general policy of urban planning. Other noted examples are the RTB TransMilenio, which has been developed since 2000 in Bogota (Colombia). The innovations of this system have made it the most solid RTB of the world and have led it to develop an extension plan of this system to seven additional cities. In Mexico City (Mexico), an RTB system has been implemented, named Metrobús, as a complement to the extensive subway system of the city. In Guatemala City (Guatemala), a trans-urban system was developed in 2009 with the aim of improving efficiency and reducing contamination indices of the transport sector in the city.

As mentioned in the introductory section, energy indicators are typically used to measure energy efficiency in countries. The most commonly used indicator of energy intensity is the ratio of energy consumption to GDP in a country. Table 4 shows the value of this indicator for the transport sector of each country and presents a ranking of "energy intensity". The countries with a lower ratio of energy consumed in transport to GDP are identified according to this indicator as those that are the most energy efficient. The table also shows the mean efficiencies that are obtained for each country with a frontier demand such as the estimated demand.[20] The correlation coefficients of both measures for each country are in some cases, such as the Dominican Republic (-0.982) and Trinidad and Tobago (-0.986), quite high and negative. This result indicates, as expected, that energy efficiency improvements are associated with a drop-offs in the energy intensity indicators. Although the correlation of these measures is high, on average, it is low in some countries (such as Brazil and Colombia). Furthermore, it is positive in two countries (Chile and Venezuela), indicating that the evolution of energy intensity indicators is associated with circumstances other than energy efficiency. Using the Spearman's rank correlation coefficient, we observe that although the rankings that are obtained by alternatively applying the criterion of energy intensity and efficiency of an energy demand model can differ (for example, Barbados and Trinidad and Tobago fall by 10 places, and Panama moves from 15 to 3 when estimating a frontier model), on average these rankings are similar, with an approximately 70% correlation between them. In summary, these results seem to confirm that the efficiency measures that are derived from the estimation of energy demand frontier models are more appropriate than those that are provided by energy intensity indicators.

[Insert Table 4 here]

5. Conclusions

In this paper, we estimate stochastic frontier demand functions to measure the level of energy efficiency of the transport sector in Latin America and the Caribbean by using panel data from 24 countries for the 1990-2010 period. The adopted approach constitutes a novel contribution to energy demand studies of the sector in this region, conferring great importance to the presented results.

Due to the different types of energy that are used in the transport sector, it is necessary to employ an index that aggregates the set of energy prices for the estimation of these demands. International energy agencies do not provide a price index for all of the countries in the sample. Thus, we construct a transitive multilateral index, which allows for consistent comparisons of energy price among countries throughout time. The construction of this price index is a relevant issue often avoided in these studies.

The estimated models are a basic stochastic frontier one and diverse latent class models that lead to obtaining differentiated demands. These models allow us to identify the reference countries in an international comparison of energy efficiency. The results indicate that the specification that best fits an energy demand is a model in which three classes are estimated using income per capita and population density as class-identifying variables. In this model, important differences in income and price elasticities can be observed. Specifically, countries with higher income per capita and

[20] For the comparisons with the rankings that are obtained based on energy intensity to make sense, the efficiency values that are shown in this table are obtained using the ALS model, as this is the only estimated model that assumes the existence of a single frontier.

lower population density have a higher probability of having a more inelastic demand in terms of price.

The estimation of the latent class model allows us to identify countries that have successfully implemented programs of improved public transport in some of their cities. The remaining countries of each class should follow the example of these countries and perform the extension or adaptation of the national transport sector policies implemented in the most efficient areas of the region, with the aim of improving energy efficiency and reducing the levels of urban contamination. Furthermore, general improvements in fuel efficiency and the transfer from private vehicle use to public transport ought to be additionally considered.

On the other hand, this paper shows that the commonly used indicators of energy intensity cannot consistently be used as a reasonable reference for energy efficiency in the transport sector. Using efficiencies that are obtained through the frontier approach, we find that although the mean efficiency is relatively high, there is a margin for energy consumption savings and, thus, for a reduction of greenhouse gas emissions. Some measures that can be adapted for this purpose are as follows: correctly assign energy prices, plan the infrastructure and land use jointly to minimize distances, balance the modal distribution, establish fiscal incentives for the use of lower consumption engines, develop fuels with reduced levels of carbon and implement awareness programs that focus on the transformation of transport use toward rational and environmentally sustainable habits.

Finally, according to the “Jevons Paradox”, it is possible that increases in energy efficiency do not involve a reduction in energy consumption and hence the energy savings predicted in the current model are not possible to reach. That situation, also called “back-fire”, is a particular case of the phenomenon known as “rebound effect”. This concept states that part of the savings from increases in the efficiency level in the use of energy can be offset by increases in the demand for energy services derived from the marginal cost reduction of those energy services. In other words, the increased efficiency in the use of a resource does not necessarily indicate a directly proportional decrease of total consumption. The rebound effect concept should be considered in future research that uses frontier approaches for the estimation of energy demands.

References

Al-Ghandoor, A., Samhouri, M., Al-Hinti, I., Jaber, J. and Al-Rawashdeh, M. (2012), "Projection of future transport energy demand of Jordan using adaptive neuro-fuzzy technique", *Energy*, 38(1), 128-135.

Ang, B.W. (2006), "Monitoring changes in economy-wide energy efficiency: From energy-GDP ratio to composite efficiency index", *Energy Policy*, 34, 574-582.

Aigner, D., Lovell, K. and Schmidt, P. (1977), "Formulation and estimation of stochastic frontier production function models", *Journal of Econometrics*, 6, 21-37.

Bauer, M., Mar, E. and Elizalde, A. (2003), "Transport and energy demand in Mexico: the personal income shock", *Energy Policy*, 31(14), 1475-1480.

Barros, C.P. and Prieto-Rodríguez, J. (2008), "A revenue-neutral tax reform to increase demand for public transport services", *Transportation Research Part A*, 42, 659-672.

Caudill, S.B., and Ford, J.M. (1993), "Biases in frontier estimation due to heteroscedasticity", *Economic Letters*, 41, 17–20.

Caudill, S.B., Ford, J.M. and Gropper, D.M. (1995), "Frontier estimation and firm-specific inefficiency measures in the presence of heteroscedasticity", *Journal of Business & Economic Statistics*, 13, 105-11.

Caves, D.W., Christensen, L.R. and Diewert, W.E. (1982), "The economic theory of index numbers and the measurement of input, output, and productivity", *Econometrica*, 50(6), 1393-1414.

ECLAC (2005), *The millennium development goals: A Latin American and Caribbean perspective*, United Nations Publications, Santiago de Chile.

ECLAC (2010), *Sustainable development in Latin America and the Caribbean: Trends, progress, and challenges in sustainable consumption and production, mining, transport, chemicals and waste management*, United Nations, Publications, Santiago de Chile.

Ceylan, H., Ceylan, H., Haldenbilen, S. and Baskan, O. (2008), "Transport energy modeling with meta-heuristic harmony search algorithm, an application to Turkey", *Energy Policy*, 36(7), 2527-2535.

Chavez-Baeza, C. and Sheinbaum-Pardo, C. (2014), "Sustainable passenger road transport scenarios to reduce fuel consumption, air pollutants and GHG (greenhouse gas) emissions in the Mexico City Metropolitan Area", *Energy*, 66(1), 624-634.

Coelli, T.J., Rao, D.S.P., O'Donnell, C.J. and Battese, G.E, (2005), *An introduction to efficiency and productivity analysis*, 2nd ed., Springer, New York.

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

Dreher, M., Wietschel, M., Göbelt, M. and Rentz, O. (1999), "Energy price elasticities of energy service demand for passenger traffic in the Federal Republic of Germany", *Energy*, 24(2), 133-140.

Elteto, O. and Koves, P. (1964), “On a problem of index number computation relating to international comparison”, *Statisztikai Szemle*, 42, 507-18.

Farsi, M., Filippini, M. and Kuenzle, M. (2005), “Unobserved heterogeneity in stochastic frontier models: An application to Swiss nursing homes”, *Applied Economics*, 37, 2127-2141.

Filippini, M. and Hunt, L.C. (2011), “Energy demand and energy efficiency in the OECD countries: A stochastic demand frontier approach”, *The Energy Journal*, 32(2), 59-80.

Filippini, M. and Hunt, L.C. (2012), “US residential energy demand and energy efficiency: A stochastic demand frontier approach”, *Energy Economics*, 34(5), 1484-1491.

Fonseca, J.R.S. and Cardoso, M.G.M.S. (2007), “Mixture-model cluster analysis using information theoretical criteria”, *Intelligent Data Analysis*, 11(2), 155-173.

Galindo, L.M. (2005), “Short-and long-run demand for energy in Mexico: a cointegration approach”, *Energy Policy* 33(9), 1179-85.

Graham, D. and Glaister, S. (2002), “The demand for automobile fuel: a survey of elasticities”, *Journal of Transport Economics and Policy*, 36, 1-26.

Greene, D.L., Kahn, J.R. and Gibson, R.C. (1999), “Fuel economy rebound effect for US household vehicles”, *The Energy Journal*, 20(3), 1-31.

Greene, W. (2004), “Distinguishing between heterogeneity and inefficiency: stochastic frontier analysis of the World Health Organization's panel data on national health care systems”, *Health Economics*, 13, 959-980.

Greene, W. (2005a), “Fixed and random effects in stochastic frontier models”, *Journal of Productivity Analysis*, 23, 7-32.

Greene, W. (2005b), “Reconsidering heterogeneity in panel data estimator of the stochastic frontier model”, *Journal of Econometrics*, 126, 269-303.

Haldenbilen, S. and Ceylan, H. (2005), “Genetic algorithm approach to estimate transport energy demand in Turkey”, *Energy Policy*, 33(1), 89-98.

Heston, A., Summers, R. and Aten, B. (2012), *Penn World Table Version 7.1*, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Hymel, K.M., Small, K.A. and Van Dender, K. (2010), “Induced demand and rebound effects in road transport”, *Transportation Research Part B: Methodological*, 44(10), 1220-1241.

Hao, H., Wang, H. and Yi, R. (2011), “Hybrid modeling of China’s vehicle ownership and projection through 2050”, *Energy*, 36(2), 1351-1361.

IEA (2014), *Energy efficiency indicators: Essentials for policy making*, International, IEA Publishing, Paris

Islas, J., Manzini, F. and Masera, O. (2007), “A prospective study of bioenergy use in Mexico”, *Energy*, 32(12), 2306-2320.

Jondrow, J., Lovell, C.A.K., Materov, I.S. and Schmidt, P. (1982), “On the estimation of technical efficiency in the stochastic frontier production function model”, *Journal of Econometrics*, 19, 233-238.

Limanond, T., Jomnonkwao, S. and Srikaew, A. (2011), "Projection of future transport energy demand of Thailand", *Energy Policy*, 39, 2754-2763.

Lu, I.J., Lin, S.J. and Lewis, C. (2008), "Grey relation analysis of motor vehicular energy consumption in Taiwan", *Energy Policy*, 36(7), 2556-2561.

Lu, I.J, Lewis, C. and Lin, S.J. (2009), "The forecast of motor vehicle, energy demand and CO_2 emission from Taiwan's road transportation sector", *Energy Policy*, 37(8), 2956-2961.

Manzini, F. (2006), "Inserting renewable fuels and technologies for transport in Mexico City Metropolitan Area", *Hydrogen Energy*, 31, 327-335.

Mundlak, Y. (1978), "On the pooling of time series and cross section data", *Econometrica*, 64, 69-85.

Murat, Y.S. and Ceylan, H. (2006), "Use of artificial neural networks for transport energy demand modeling", *Energy Policy*, 34(17), 3165-3172.

Orea, L. and Kumbhakar, S. (2004), "Efficiency measurement using stochastic frontier latent class model", *Empirical Economics*, 29, 169-183.

Pradhan, S., Bahadur, B. and Bhusan, V. (2006), "Mitigation potential of greenhouse gas emission and implications on fuel consumption due to clean energy vehicles as public passenger transport in Kathmandu Valley of Nepal: a case study of trolley buses in ring road", *Energy*, 31(12), 1748-1760.

Reifschneider, D. and Stevenson, R. (1991), "Systematic departures from the frontier: A framework for the analysis of firm inefficiency", *International Economic Review*, 32: 715-723.

Sa'ad, S. (2010), "Improved technical efficiency and exogenous factors in transportation demand for energy: an application of structural time series analysis to South Korean data", *Energy*, 35(7), 2745-2751.

Samini, R. (2003), "Road transport energy demand modelling. A cointegration approach", *Energy Economics*, 17(4), 329-339.

Schmidt, P. and Sickles, R.C. (1984), "Production frontiers and panel data", *Journal of Business and Economic Statistics*, 2(4), 367-374.

Small, K.A. and Van Dender, K. (2007), "Fuel efficiency and motor vehicle travel: the declining rebound effect", *The Energy Journal*, 28(1), 25–51.

Stead, D. (2001), "Transport intensity in Europe – indicators and trends", *Transport Policy*, 8, 29-46.

Suganthi, L. and Samuel, A.A. (2012), "Energy models for demand forecasting–A review", *Renewable and Sustainable Energy Reviews*, 16(2), 1223-1240.

Szulc, B. (1964), "Indices for multiregional comparisons", *Przeglad Statystyezny 3, Statistical Review 3*, 239-54.

Wohlgemuth, N. (1997), "World transport energy demand modeling: methodology and elasticities", *Energy Policy*, 25 (14–15), 1109-1119.

Zhang, M., Mu, H., Li, G. and Ning, Y. (2009), "Forecasting the transport energy demand based on PLSR method in China", *Energy*, 34(9), 1396-1400.

Table 1. Descriptive statistics

Variable	*Units*	*Mean*	*Std. Dev.*	*Max.*	*Min.*
Q	Thousands of toe	6,141	12,434	69,384	18
Y	Millions of US Dollars (2005)	164,968	339,168	1,800,000	713
POP	Thousands of inhabitants	20,517	38,114	195,498	91
P	Index	174.56	108.64	850.66	3.76
ST	%	4.02	1.59	12.74	1.07
DEN	Thousands of inhabitants / km^2	0.10	0.14	0.63	0.00

Table 2. Standard frontier demand model

	ALS model		
Variable	*Coeff.*		*t-ratio*
Intercept	7.098	***	405.450
ln (Y)	0.810	***	39.720
ln (POP)	0.182	***	8.834
ln (P)	-0.229	***	-15.138
ST	0.047	***	7.103
ln (DEN)	-0.096	***	-12.031
t	0.013	***	6.960
½ t^2	-0.001		-1.537
$\sigma = (\sigma_v^2+\sigma_u^2)^{(1/2)}$	0.257	***	590.578
$\lambda = \sigma_u/\sigma_v$	0.886	***	7.411
σ_v	0.192		
σ_u	0.170		
Log LF		52.689	

Significance: * p<0.1, ** p<0.05, *** p<0.01

Table 3. Frontier demands with latent class including separating variables

	LCSFM with two classes						LCSFM with three classes								
	Class 1			Class 2			Class 1			Class 2			Class 3		
Variable	***Coeff.***		***t-ratio***	***Coeff.***		***t-ratio***	***Coeff.***		***t-ratio***	***Coeff.***		***t-ratio***	***Coeff.***		***t-ratio***
Intercept	7.180	***	271.915	6.903	***	371.389	7.367	***	195.898	7.091	***	192.733	6.894	***	522.491
ln (Y)	0.784	***	24.069	0.637	***	35.489	0.566	***	22.754	0.179	***	3.687	0.649	***	35.921
ln (POP)	0.188	***	5.491	0.280	***	17.498	0.431	***	16.176	0.742	***	16.154	0.247	***	11.754
ln (P)	-0.188	***	-17.045	-0.175	***	-5.417	-0.161	***	-13.869	-0.288	***	-13.057	-0.407	***	-10.379
ST	0.094	***	8.745	0.037	***	7.048	0.044	***	4.966	0.002		0.307	-0.008		-0.811
ln (DEN)	-0.067	***	-5.137	-0.046	***	-6.349	0.007		0.714	0.125	***	6.868	-0.030	***	-4.016
t	0.006	***	2.774	0.016	***	8.642	0.009	***	4.985	0.042	***	16.986	0.030	***	12.527
½ t^2	0.000		0.426	-0.003	***	-5.835	0.000		-0.505	0.003	***	5.718	-0.001	***	-2.720
$\sigma = (\sigma_v^2+\sigma_u^2)^{(1/2)}$	0.246	***	12.009	0.171	***	11.022	0.166	***	7.336	0.112	***	5.391	0.135	***	12.703
$\lambda = \sigma_u/\sigma_v$	2.961	***	2.974	2.320	***	3.147	1.003	*	1.932	0.999		1.401	3.040	***	3.937
σ_v	0.079			0.068			0.117			0.079			0.042		
σ_u	0.233			0.157			0.118			0.079			0.128		
Class membership probabilities															
Intercept	0.236		0.507	-	-	-	1.036		1.237	0.736		0.870	-	-	-
ln (Y/POP)	0.639		0.674	-	-	-	4.293	**	2.250	3.082	*	1.770	-	-	-
ln (DEN)	-0.543		-1.397	-	-	-	-2.090	**	-2.485	-1.005		-1.387	-	-	-
Prior Prob.		0.559			0.441			0.477			0.354			0.169	
Log LF			281.782								398.356				

Significance: * p<0.1, ** p<0.05, *** p<0.01

Table 4. Country ranking using energy intensity and energy efficiency

Country	Indicator (Energy/GDP)		Frontier demand		Correlation *(EI Vs Eff.)*
	EI	*Ranking*	*Eff.*	*Ranking*	
Argentina	0.037	14	0.845	19	-0.897
Barbados	0.019	1	0.885	11	-0.938
Bolivia	0.042	19	0.869	15	-0.883
Brazil	0.034	12	0.872	14	-0.241
Chile	0.044	20	0.844	20	0.161
Colombia	0.032	8	0.896	7	-0.061
Costa Rica	0.032	9	0.875	13	-0.720
Ecuador	0.055	22	0.828	22	-0.962
El Salvador	0.026	5	0.902	5	-0.931
Granada	0.029	7	0.877	12	-0.807
Guatemala	0.024	2	0.910	4	-0.952
Guyana	0.066	24	0.846	18	-0.956
Honduras	0.033	10	0.890	9	-0.925
Jamaica	0.038	16	0.813	24	-0.914
Mexico	0.040	18	0.861	17	-0.814
Nicaragua	0.040	17	0.888	10	-0.946
Panama	0.037	15	0.914	3	-0.893
Paraguay	0.054	21	0.815	23	-0.951
Peru	0.025	3	0.933	1	-0.763
Dominican Rep.	0.026	4	0.898	6	-0.982
Suriname	0.035	13	0.891	8	-0.906
Trinidad and Tobago	0.033	11	0.834	21	-0.986
Uruguay	0.028	6	0.924	2	-0.706
Venezuela	0.062	23	0.868	16	0.153
Spearman's rank correlation coefficient between both rankings					0.701

Note: *EI* stands for *Energy Intensity* and *Eff.* is the abbreviation of *Efficiency*

Figure 1. Final energy consumption by activity sector (average for Latin America and the Caribbean in the 1990-2010 period)

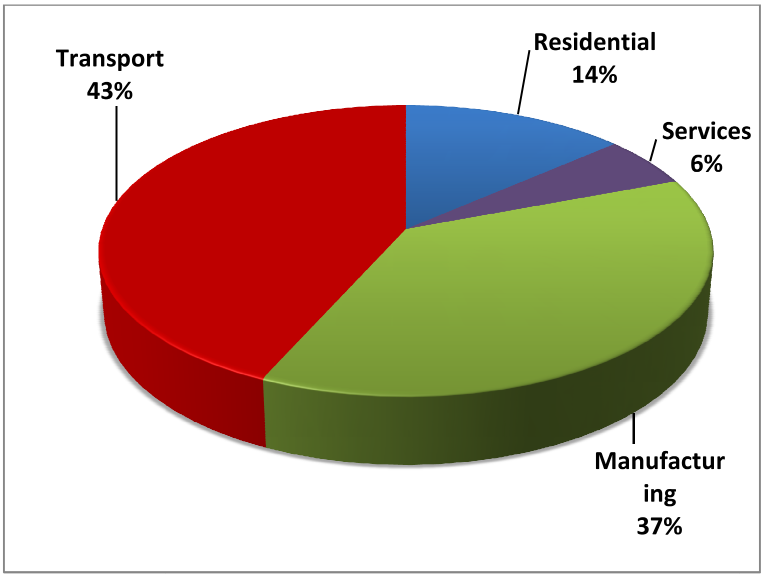


Figure 2. Energy consumption in tons of oil equivalent (toe) per capita in transport (average for Latin America and the Caribbean in the 1990-2010 period)

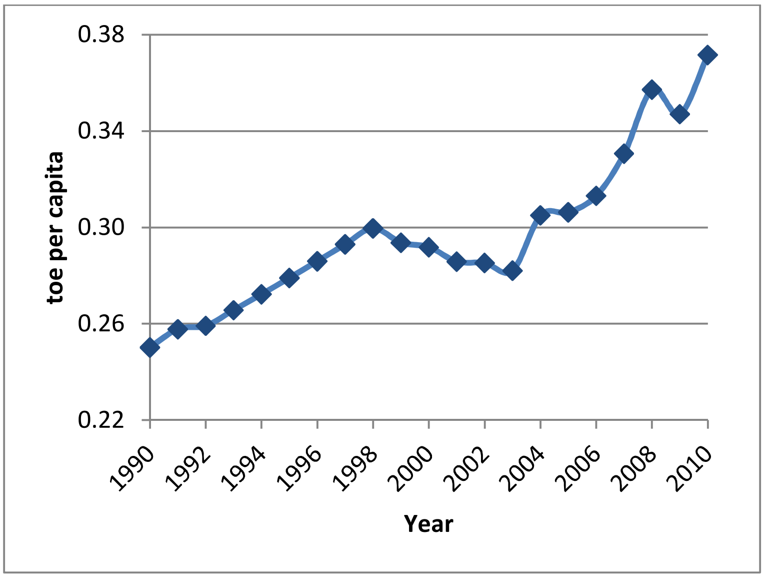


Figure 3. Price index for energy in the transport sector (average for Latin America and the Caribbean in the 1990-2010 period)

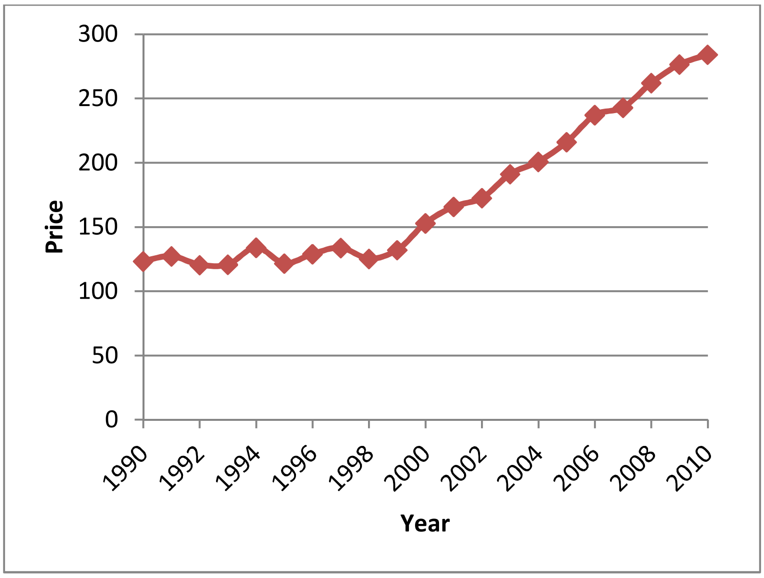


Figure 4. Approaches in the estimation of energy demand functions

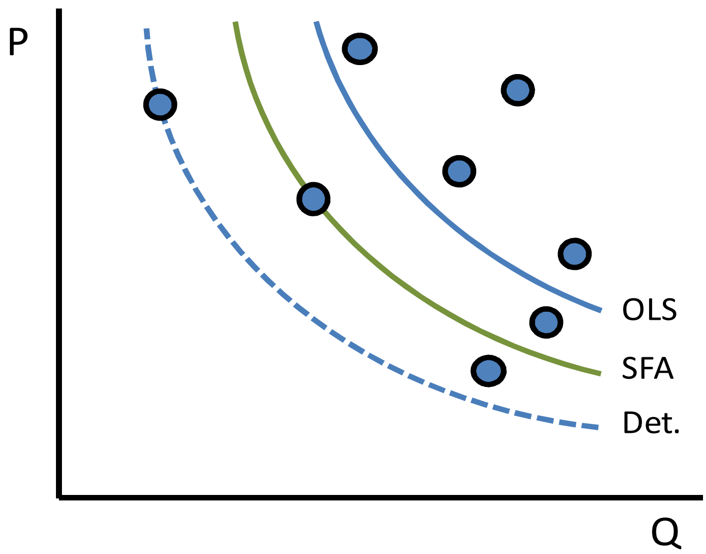


Figure 5. Decomposition of the random error term in a stochastic frontier demand

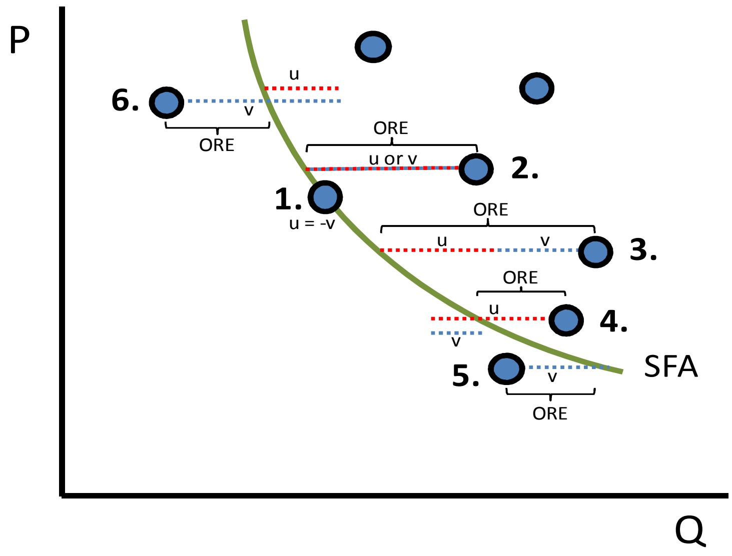
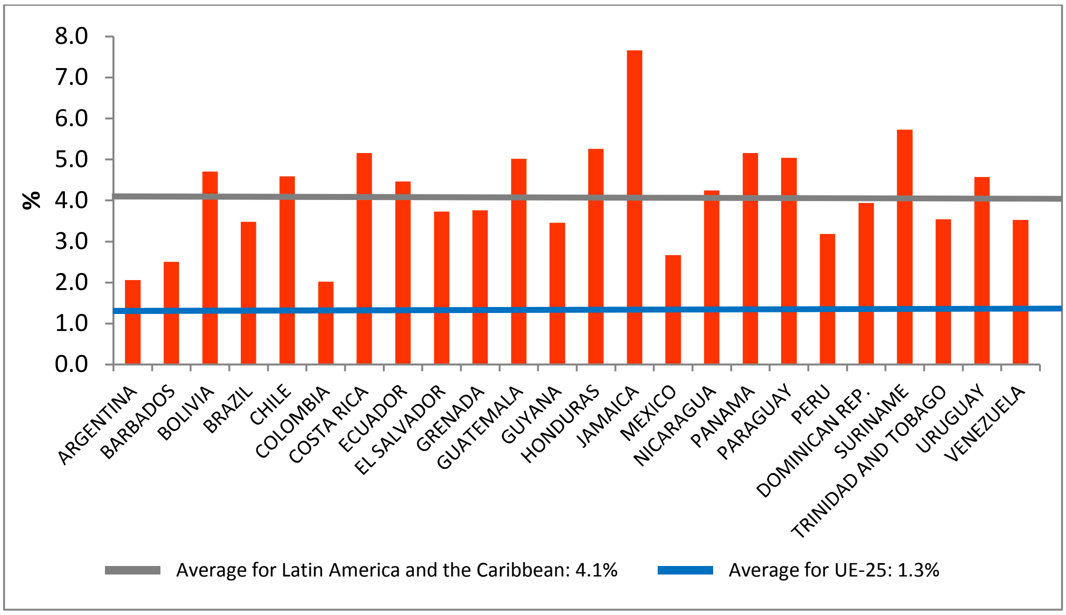


Figure 6. Average annual growth rate of energy consumption in transport for Latin America and the Caribbean, 1990-2010 (percentage)



Sources: ECLAC and EUROSTAT

Figure 7. Transitive multilateral price index of energy in the transport sector for Latin America and the Caribbean, 1990-2010

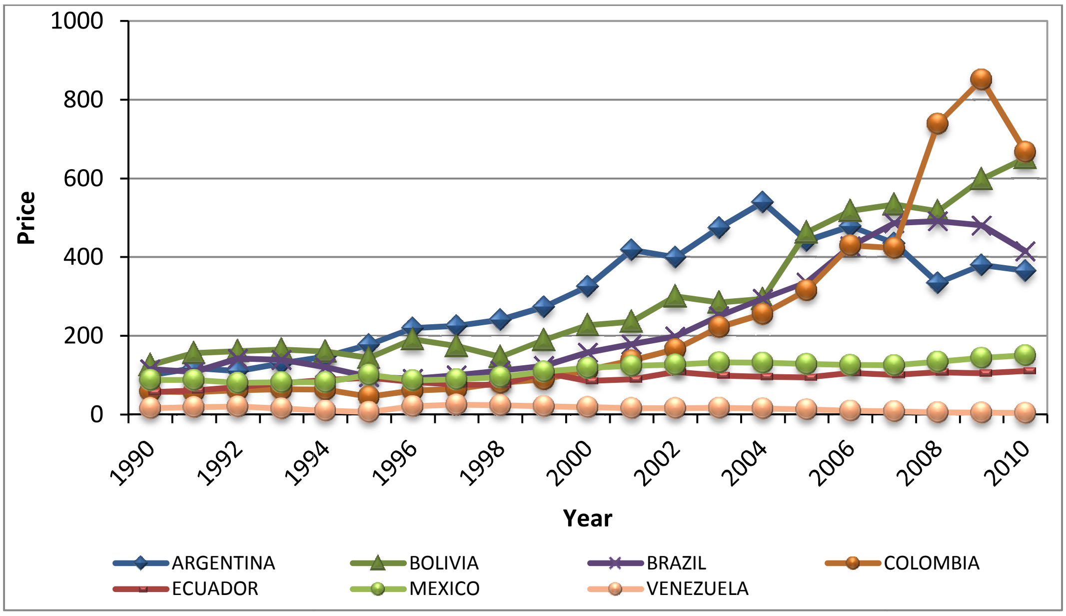


Figure 8. Model selection tests

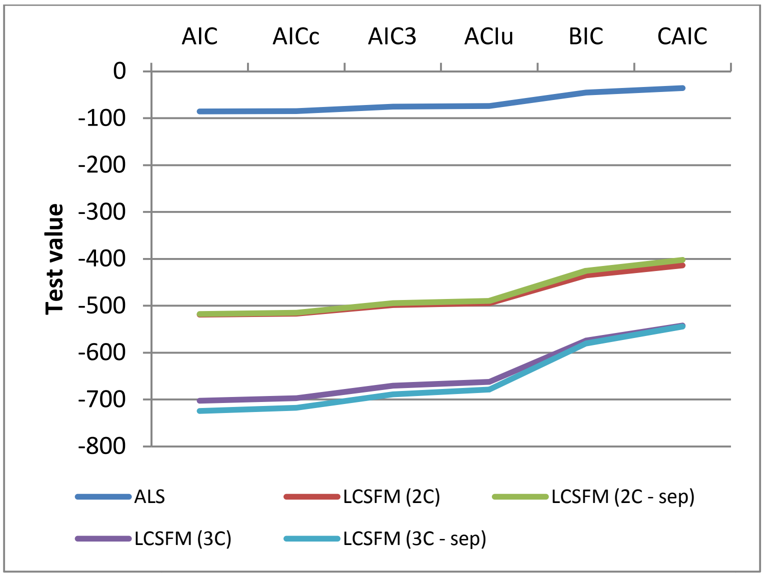
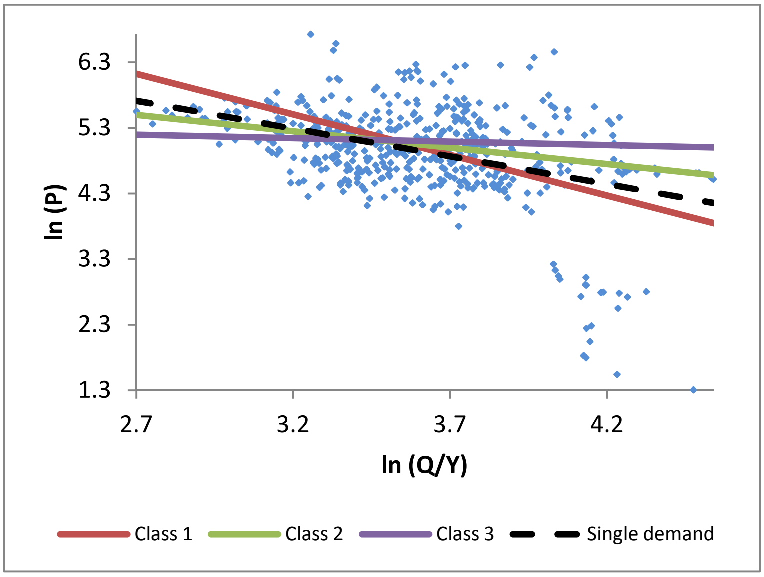


Figure 9. Linear demands obtained on the basis of observed values



APPENDIX

Construction of the price index

The OLADE provides information on the prices and quantities consumed of the different types of energy that are used in the transport sector of Latin America and the Caribbean. The categories that appear in their database are as follows: natural gas, liquid gas, electricity, various types of gasoline, kerosene, diesel oil and fuel oil. However, this agency does not provide a general price of energy for these countries. Thus, to estimate aggregate energy demand in transport, it is necessary to obtain an indicator or index that accounts for the distinct components in the energy consumption of the sector. In general, a compound price index can be defined as follows:

$$PI_{0t} = \frac{\sum_{m=1}^{M} p_{mt} q_{mt}}{\sum_{m=1}^{M} p_{m0} q_{m0}} \qquad \text{(A1)}$$

where PI_{0t} measures the change in value of the total of the M energy components between the base period 0 and final period t. In this type of index, it is difficult to distinguish between the changes that only occur in prices and the change in consumed quantities. The two indices that are most commonly used in practice and calculated by international agencies for total energy consumption, such as those calculated by the IEA, are Laspeyres and Paasche. In the former, the quantities that are consumed in the base year (q_{m0}) are used as weights both in the numerator and in the denominator. Thus, this index isolates the change in prices without accounting for changes in consumption patterns. The second type of index uses energy quantities from the current period (q_{mt}) as weights, thus simultaneously including variations in prices and quantities. These two indices, therefore, represent two extreme cases and only coincide when relative prices do not experience any variation (i.e., p_{mt}/p_{m0} is constant).

However, there are alternatives that combine both approaches to address this issue, such as the Fisher and Törnqvist indices. Nevertheless, all of these indices present the same problem. Specifically, they allow for comparisons of a country with itself throughout time and comparisons between countries measured in price changes (if the same base year is imposed for all countries in the sample), but they do not allow for comparisons of price levels between countries throughout time.

Studies that use international data must employ an index that overcomes this difficulty. The solution to this problem involves obtaining transitive multilateral comparisons (as named in the literature on index numbers) between countries, as proposed by Elteto and Koves (1964) and Szulc (1964). This method, known as EKS, was used by Caves *et al.* (1982) to obtain transitive Törnqvist indices. The formula, in line with Coelli *et al.* (2005), is as follows:

$$\ln PI_{ij}^{CCD} = \frac{1}{2}\sum_{m=1}^{M}\left(\omega_{mj} + \bar{\omega}_m\right)\left(\ln p_{mj} - \overline{\ln p_m}\right) - \frac{1}{2}\sum_{m=1}^{M}\left(\omega_{mi} + \bar{\omega}_m\right)\left(\ln p_{mi} - \overline{\ln p_m}\right) \qquad \text{(A2)}$$

where ω_{mi} represents the importance held by component m in the energy expenditure of the transport sector of the country i and $\bar{\omega}_m$ is the arithmetic mean of these expenditure amounts. Furthermore, $\overline{\ln p_m}$ represents the average price of the energy component m for the set of countries.

The intuitive interpretation of equation (A2) is that to compare the price indices of two countries, each of them is compared to the average country and then the differences from this mean are calculated. Logically, as opposed to other indices, when an observation is added or subtracted from the sample, all values should be recalculated due to changes in the mean of the sample.

It should be mentioned that in the current empirical application, the use of an approach such as the proposed by Caves *et al.* (1982) in the construction of the price index significantly improves the quality of fitting that is obtained when estimating the models. If a Paasche- or Laspeyres-type index is used rather than a transitive multilateral index, the logarithm of the likelihood function falls sharply and achieves negative values. The use of these simpler indices in practice implies the assumption that each country has a specific individual effect. In this case, we artificially introduce heterogeneity into the model. Thus, the model must be estimated by including individual effects, as in the TFE and TRE models.