Centre for Energy Policy and Economics
Swiss Federal Institutes of Technology

US Residential Energy Demand and Energy Efficiency: A Stochastic Demand Frontier Approach

Massimo Filippini
Lester Hunt

CEPE Working Paper No. 83
April 2012

CEPE
Zurichbergstrasse 18 (ZUE E)
CH-8032 Zurich
www.cepe.ethz.ch

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

US Residential Energy Demand and Energy Efficiency: A Stochastic Demand Frontier Approach*

Massimo **Filippini** and Lester C **Hunt**

Centre for Energy Policy and Economics (cepe), ETH Zurich
and
Department of Economics, University of Lugano, Switzerland

Massimo.Filippini@usi.ch

Surrey Energy Economics Centre (SEEC) and Research Group on Lifestyles Values and Environment (RESOLVE), Department of Economics, University of Surrey, UK

L.Hunt@surrey.ac.uk

Abstract

This paper estimates a US frontier residential aggregate energy demand function using panel data for 48 'states' over the period 1995 to 2007 using stochastic frontier analysis (SFA). Utilizing an econometric energy demand model, the (in)efficiency of each state is modelled and it is argued that this represents a measure of the inefficient use of residential energy in each state (i.e. 'waste energy'). This underlying efficiency for the US is therefore observed for each state as well as the relative efficiency across the states. Moreover, the analysis suggests that energy intensity is not necessarily a good indicator of energy efficiency, whereas by controlling for a range of economic and other factors, the measure of energy efficiency obtained via this approach is. This is a novel approach to model residential energy demand and efficiency and it is arguably particularly relevant given current US energy policy discussions related to energy efficiency.

JEL Classification: D2, Q4.

Keywords: US residential energy demand; efficiency and frontier analysis; state energy efficiency.

* This is a revised version of the paper previously published as SEEDS Working Paper No. 130, November 2010 (http://www.seec.surrey.ac.uk/research/SEEDS/SEEDS130.pdf).

Acknowledgements
A preliminary version of the paper was presented at the 33rd IAEE International Conference, Rio de Janeiro, Brazil, 2010; the 3rd International Workshop on Empirical Methods in Energy Economics, SEEC, University of Surrey, UK, 2010; the 11th IAEE European Conference, Vilnius, Lithuania, 2010 and The University of Maryland, 2011 and we are grateful to participants for their comments and suggestions.

1 Introduction

The promotion of energy efficiency policy is seen as a very important activity by both the International Energy Agency (IEA) and the Energy Information Administration (EIA) (e.g. see IEA, 2009). Moreover, the role of energy efficiency in reducing energy consumption and emissions remains a key policy objective for governments across the globe; and the US is no exception. Since the beginning of the Obama administration, there have been many policy announcements involving energy efficiency in one way or another; you just have to look at the US Department of Energy press web site to see the many different announcements.[1] Nevertheless, it is worth noting that a number of the announcements during the Obama period build upon initiatives from the Bush administration, such as The Energy Efficiency and Conservation Block Grant (EECCBG); initially put in place in 2007 to help implement energy efficiency and conservation measures.[2] Given its importance, and the many millions of US dollars allocated across the different states it is vital that US policy makers understand, and are able to clearly measure, the relative energy efficiency across the different states. However, generally this is not the case, which is not a new problem; the EIA (1995) report states:

> "Energy efficiency is a vital component of the Nation's energy strategy. One of the Department of Energy's missions are to promote energy efficiency to help the Nation manage its energy resources. *The ability to define and measure energy efficiency is essential to this objective*. In the absence of consistent defensible measures, energy efficiency is a vague, subjective concept that engenders directionless speculation and confusion rather than insightful analysis. ... The task of defining and measuring energy efficiency and creating statistical measures as descriptors is a daunting one." (p. vii, *our emphasis*).

This supports the view above, but the EIA (1995) report goes on to discuss the use of energy intensity as a "measurement indicator of energy efficiency" (p. vii) highlighting that

[1] www.energy.gov/news/releases.htm.

[2] www1.eere.energy.gov/wip/eecbg.html.

energy intensity and energy efficiency are often used interchangeably; furthermore, energy intensity might not reflect certain factors that would allow energy intensity to approximate energy efficiency accurately. In particular, trends in different measures of energy intensity are generally suggestive of trends in energy efficiency but the trends in energy intensity are likely to be influenced by factors other than just energy efficiency. Moreover, the EIA (1995) report states that

> "it is virtually impossible to remove, or even to consider, all of the behavioral or structural factors that would be necessary to obtain a pure measurement of energy efficiency, however broadly energy efficiency may be defined." (p. vii).

This clearly highlights the problems in trying to measure energy efficiency in general and the use of energy intensity in particular as a proxy for it. Furthermore, given the problems with energy intensity, it shows that there is a need to 'control' for other important factors in order to get a 'pure' measure of energy efficiency. This, therefore, is one of the key aims of this paper with respect to the US residential sector.

The EIA (1995) report goes on to consider the measurement of energy intensity in a number of sectors of the US economy attempting, where possible, to remove the influence of such factors as weather, capacity, and inventory changes that are commonly viewed as not related to changes in energy efficiency. For the residential sector, the EIA (1995) report suggests four energy intensity measures applicable as proxies for energy efficiency: i) million BTUs per building; ii) million BTUs per household; iii) thousand BTUs per square foot; and iv) million BTUs per household member.[3] However, the report suggests that these are imperfect and that "No single energy-intensity indicator for the residential sector stands

[3] BTU = British Thermal Unit; the quantity of heat required to raise the temperature of 1 pound of liquid water by 1 degree Fahrenheit at the temperature at which water has its greatest density.

out as clearly superior to the others. The choice of indicator depends on the questions asked and on data and resource availability" (p. 16).

Some approaches have been proposed in the energy economics literature in order to overcome the problems of some of these simple efficiency indicators; such as Index Decomposition Analysis (IDA) and Frontier Analysis (FA). IDA is basically a bottom-up framework used to create energy efficiency indicators.[4] For instance, the US Department of Energy has introduced an Energy Intensive Index using the decomposition approach that attempts to separate the difference factors that affect energy efficiency from non-efficiency factors.[5] Whereas FA is based on the estimation of a parametric, as well as a non-parametric, best practice frontier for the use of energy where the level of energy efficiency is computed as the difference between the actual energy use and the predicted energy use.

Huntington (1994) discusses the relationship between energy efficiency and productive efficiency using the production theory framework. Zhou and Ang (2008) is an example of a non-parametric approach, where the energy efficiency performance of 21 OECD countries over 5 years (1997-2001) is measured using Data Envelopment Analysis (DEA).[6] Examples of the use of parametric FA at the sectoral level are Buck and Young (2007) who measured the level of energy efficiency of a sample of Canadian commercial buildings and Boyd (2008) who estimated an energy use frontier function for a sample of wet corn milling plants. In addition, Filippini and Hunt (2011) estimate a panel frontier

---

[4] See Boyd and Roop (2004) and Ang (2006) for a general discussion and application of this method.

[5] See www1.eere.energy.gov/ba/pba/intensityindicators/. It is argued that the new index gives a more accurate representation of intensity change associated with energy efficiency improvement than the simple energy/activity ratios.

[6] For a more general discussion on the use of DEA in energy analysis, see Zhou et al. (2008).

whole economy aggregate energy demand function for 29 OECD countries over the period 1978 to 2006 using parametric Stochastic Frontier Analysis (SFA).

As stated above, the aim of this paper is to attempt to construct and measure the 'underlying energy efficiency' for the US residential sector across 48 'states',[7] building on previous work by Filippini and Hunt (2011). This draws upon different strands of the energy economics research literature; in particular, frontier estimation and energy demand modelling. An aggregate energy demand frontier function is estimated in order to isolate the measure of 'underlying energy efficiency'; explicitly controlling for income and price effects, population, household size, weather, types of housing, regional effects, and a common Underling Energy Demand Trend (the UEDT, capturing both 'exogenous' technical progress and other exogenous factors[8]). Furthermore, the UEDT needs to be specified in such a way that it is 'non-linear' and therefore could increase and/or decrease over the estimation period,[9] and given a panel data set is used, this is achieved by the inclusion of time dummies.[10]

In summary, in order to try to uncover these different influences, a general energy demand relationship for US residential energy demand relating energy consumption to economic activity and the real energy price is estimated for a panel of 48 states; but controlling for other important factors that vary across states and hence can affect a states' residential energy demand. This model attempts to isolate the 'underlying energy efficiency' for each state, defined with respect to a benchmark, e.g. a best practice state in

[7] The reason for the use of only 48 states is explained below.

[8] Hence, this method allows for the impact of 'endogenous' technical progress' through the price effect and 'exogenous' technical progress through the UEDT.

[9] As advocated by Hunt et al. (2003a and 2003b)

[10] As proposed by Griffin and Schulman (2005) and Adeyemi and Hunt (2007).

the use of residential energy by estimating a ‘common energy demand’ function across states, with homogenous income and price elasticities, and responses to other factors, plus a homogenous UEDT. This is seen as important, given the need to isolate the ‘underlying energy efficiency’ across the different states.[11] Consequently, once these effects are controlled for, it allows for the estimation of the ‘underlying energy efficiency’ for each state and the differences across the panel of states.

The paper is organized as follows. The next section, discusses the rationale and specification of the energy demand frontier function, with the data and econometric specification introduced in Section 3. The results of the estimation are presented in Section 4, with a summary and conclusion in the final section.

2 An aggregate frontier energy demand model

Residential demand for energy is derived from the demand for a warm house, lighting, cooked food, hot water, etc., and can be specified using the basic framework of household production theory. According to this theory, households purchase market ‘goods’ that serve as inputs in the production processes, to produce the ‘commodities’ which appear as arguments in the household's utility function. Within the framework of the household production theory, the aggregate residential energy demand is an input demand function.[12]

[11] The UEDT implicitly includes exogenous technical progress of the appliance and building stock and it could be argued that even though technologies are available to each state they are not necessarily installed at the same rate. However, it is assumed that this results from different behaviour across states and reflects ‘inefficiency’ across states; hence, it is captured by the different (in)efficiency terms for all states.

[12] For a presentation of the household production theory, see Deaton and Muellbauer (1980). See Filippini (1999) and Banfi et al. (2005) for an application of household production theory to energy demand analysis.

Given the discussion above, it is assumed that there exists an aggregate US residential energy demand relationship for a panel of states, as follows:

$$E_{it} = E(P_{it}, Y_{it}, POP_{it}, AHS_{it}, HDD_{it}, CDD_{it}, SDH_{i}, D_{t}, EF_{it}) \quad (1)$$

where E_{it} is aggregate residential energy consumption, Y_{it} is real income, P_{it} is the real energy price, POP_{it} is population, AHS_{it} is the average household size, HDD_{it} are the heating degree days, and CDD_{it} are the cooling degree days; all for state *i* in year *t*. SDH_i is the share of detached houses for state *i*, D_t is a series of time dummy variables and EF_{it} is the level of 'underlying energy efficiency' of the US residential sector for state *i* in year *t*. The 'underlying energy efficiency' could incorporate a number of factors that will differ across states, including the different technical appliance and capital equipment, different regulations as well as different social behaviours, norms, lifestyles and values. Hence, a low level of 'underlying energy efficiency' implies an inefficient use of energy (i.e. 'waste energy'), so that in this situation, awareness of energy conservation could be increased in order to reach the 'optimal' energy demand function. Nevertheless, from an empirical perspective, when using US residential aggregate energy data, the aggregate level of energy efficiency of residential appliances is not observed directly. Therefore, this 'underlying energy efficiency' indicator has to be estimated. Consequently, in order to estimate the residential level of 'underlying energy efficiency' and identify the best practice system in term of energy utilization, the stochastic frontier function approach introduced by Aigner et al. (1977) is used.

The stochastic frontier function has generally been used in production theory to measure econometrically the economic performance of production processes. The central concept of the frontier approach is that in general the function gives the maximum or minimum level of an economic indicator attainable by an economic agent. For an input demand function the frontier gives the minimum level of input used by a firm or a household for any given level of output; hence, the difference between the observed input and the cost-minimizing input demand represents both technically as well allocative inefficiency.[13] In the case of an aggregate residential energy demand function, used here, the frontier gives the minimum level of energy consumption necessary for the residential sector to produce any given level of energy services. In principle, the aim here is to apply the frontier function concept in order to estimate the baseline energy input demand, which is the frontier that reflects the demand of the residential sector of a state that have and use high efficient equipment and production process. This frontier approach allows the possibility to identify if a state is, or is not, on the frontier. Moreover, if a state is not on the frontier, the distance from the frontier measures the level of energy consumption above the baseline demand, e.g. the level of energy inefficiency.

The approach used in this study is therefore based on the assumption that the level of the energy efficiency of the residential sector can be approximated by a one-sided non-negative term, so that a panel log-log functional form of Equation (1) adopting the SFA approach proposed by Aigner et al. (1977) can be specified as follows:

$$e_{it} = \alpha + \alpha^{p} p_{it} + \alpha^{y} y_{it} + \alpha^{pop} pop_{it} + \alpha^{ahs} ahs_{it} + \alpha^{hdd} hdd_{it} + \alpha^{cdd} cdd_{it} + \alpha^{SDH} SDH_{i} + \alpha^{t} Dt + v_{it} + u_{it} \qquad (2)$$

[13] See Kumbhakar and Lovell (2000, p. 148) for a discussion on the interpretation of the efficiency in an input demand function.

where e_{it} is the natural logarithm of aggregate energy consumption (E_{it}), p_{it} is the natural logarithm of the real price of energy (P_{it}), y_{it} is the natural logarithm of real income (Y_{it}), pop_{it} is the natural logarithm of population (POP_{it}), ahs_{it} is the natural logarithm of the average household size (AHS_{it}), hdd_{it} is the natural logarithm of the heating degree days (HDD_{it}), cdd_{it} is the natural logarithm of the cooling degree days (CDD_{it}) and SDH_i, and D_t are as defined above. Furthermore, the error term in Equation (2) is composed of two independent parts. The first part, v_{it}, is a symmetric disturbance capturing the effect of noise and, as usual, is assumed to be normally distributed. The second part, u_{it}, (which represents the level of 'underlying energy efficiency' EF_{it} in equation (1)) is interpreted as an indicator of the inefficient use of energy, e.g. the 'waste energy'.[14] It is a one-sided non-negative random disturbance term that can vary over time, assumed to follow a half-normal distribution.[15] An improvement in the energy efficiency of the equipment or in the use of energy through a new production process will increase the level of energy efficiency of a state. The impact of technological, organizational, and social innovation in the production and consumption of energy services on the energy demand is therefore captured in several ways: the time dummy variables, the indicator of energy efficiency and through the price effect.[16]

In summary, Equation (2) is estimated in order to estimate 'underlying energy efficiency' for each state in the sample. The data and the econometric specification of the estimated equations are discussed in the next section.

[14] As discussed later, some SFMs assume that the level of efficiency is constant over time (u_i).

[15] It could be argued that the half-normal distribution is a strong assumption for *EF*, but it does allow the 'identification' of the efficiency for each state separately. This is a standard assumption used in the production frontier literature; see Kumbhakar and Lovell (2000, p. 148) for a discussion.

[16] Of course, the time dummies capture a general and not a state specific UEDT.

3. Data and econometric specification

The study is based on a balanced US panel data set for a sample of 48 states ($i = 1, \ldots, 48$) over the period 1995 to 2007 ($t = 1995\text{-}2007$). For the purposes of this paper attention is restricted to the contiguous states (i.e. Alaska and Hawaii are excluded) as is Rhode Island because of incomplete information whereas the District of Columbia is included and considered as a separate 'state'. The data set is based on information taken from the US EIA database called States Energy Data System, from the US Department of Commerce, the US Census Bureau and the National Climatic Data Center at NOAA.

E_{it} is each state's aggregate residential energy consumption for each year in trillion BTUs, Y_{it} is each state's real disposable personal income for each year in thousand million US 1982$, P_{it} is each state's real energy price for each year in per million BTUs 1982$. Residential energy consumption figures and prices are provided by the EIA. Population (POP_{it}) and real disposable personal income are from the Bureau of Economic Analysis of the US Census Bureau. The heating and cooling degree days (HDD_{it} and CDD_{it}) are obtained from the National Climatic Data Center at NOAA. The average size of a household (AHS_{it}) is obtained by dividing population by the number of housing units, where the latter come from the US Census Bureau and the share of detached houses for each state (SDH_i) is based on the 2000 census obtained also from the Census Bureau. Descriptive statistics of the key variables are presented in Table 1.

Table 1: Descriptive Statistics

Variable Description	Name	Mean	Std. Dev.	Minimum	Maximum
Energy consumption *(Trillion Btu)*	***E***	227.630	209.64	19.80	915.6
Real disposable personal income *(Thousand million 1982US$)*	***Y***	588751.3	101167	6072.44	646019
Real Price of energy *(Per million Btu)*	***P***	15.29	4.20	7.35	32.50
Population *(1,000)*	***POP***	5863	6275	485	36377
Average Household size *(No. of people per housing unit)*	***AHS***	2.35	0.16	1.89	2.99
Heating degree days *(Base: 65F)*	***HDD***	5087	1998	555	10745
Cooling degree days *(Base: 65F)*	***CDD***	1142	796	128	3870
Share of detached houses	***SDH***	62.30	9.74	13.20	74

It is important to discuss the literature on the estimation of a Stochastic Frontier Model (SFM) using panel data, given the econometric specification of the model. This literature identifies at least three models that could be used in this empirical analysis, such as: i) the Pooled Model (PM hereafter) is the SFM in its original form proposed by Aigner, et al., (1977) as used in panel data; ii) the Random Effects Model (REM hereafter) proposed by Pitt and Lee (1981) who interpreted the panel data random effects as inefficiency rather than heterogeneity; and iii) the True Random Effects model (TREM hereafter) proposed more recently by Greene (2005a and 2005b).[17] A shortcoming of the REM is that any unobserved, time-invariant, group-specific heterogeneity is considered as inefficiency. Moreover, the level of efficiency does not vary over time. In this case, the inefficiency term is given by u_i and not u_{it}. In order to solve this problem using panel data, Greene (2005a and 2005b) proposed the TREM by extending the PM by adding a random

[17] Schmidt and Sickles (1984) and Battese and Coelli (1992) presented variations of this model. For a general discussion on the use of SFMs in the energy sector, see Farsi and Filippini (2009).

individual effect.[18] In the TREM the general constant term, *α,* in equation (2), is substituted with a series of state-specific random effects that take into account all unobserved socioeconomic and environmental characteristics that are time-invariant. The TREM is therefore able to distinguish time invariant unobserved heterogeneity from the time varying level of efficiency component. In this way, the TREM arguably overcomes some of the limitations of conventional frontier panel data models (see Greene, 2005a and 2005b); however, it produces efficiency estimates that do not include the persistent inefficiencies that might remain more or less constant over time. To the extent that there are certain sources of energy efficiency that result in time-invariant excess energy consumption, the estimates of these models provide relatively high levels of energy efficiency.

As discussed in Farsi et al. (2005b) all these approaches (PM, REM and TREM) can suffer from the 'unobserved variables bias', because the unobserved characteristics may not be distributed independently of the explanatory variables.[19] In order to address this econometric problem, this study follows the approach taken by Farsi et al. (2005b) by using a Mundlak version of the REM originally proposed by Pitt and Lee (1981). The Mundlak version of the REM (MREM hereafter) is based upon Mundlak's (1978) modification of the REM for the general specification; whereby the correlation of the individual specific effects (u_i) and the explanatory variables are considered in an auxiliary equation given by:

$$u_i = AX_i \pi + \gamma_i \qquad AX_i = \frac{1}{T}\sum_{t=1}^{T} X_{it}\,,\ \gamma_i \sim iid(0, \sigma_\delta^2) \tag{3}$$

[18] For a successful application of these models in network industries, see Farsi, et al. (2005a) and Farsi, et al. (2006).

[19] Of course, this heterogeneity bias can be reduced to some extent by introducing several explanatory variables and by considering a relatively long period. This approach was adopted by Filippini and Hunt (2011) in estimating an energy demand frontier model for OECD countries using a PM and in that case, the coefficients obtained using different models were relatively similar.

where X_{it} is the vector of all explanatory variables, AX_i is the vector of the averages of all the explanatory variables and π is the corresponding vector of coefficients.[20] Equation (3) is readily incorporated in the main frontier equation (2) and estimated using the REM. Nevertheless, in a frontier model the error term is a composite asymmetric term, consequently, the estimated coefficients are not the within estimators as in Mundlak's classical formulation. However, since the correlation between the individual effects and the explanatory variables is at least partially captured in the model, the heterogeneity bias is expected to be relatively low. Moreover, as shown in Farsi et al. (2005b), the application of Mundlak's adjustment to the REM frontier framework decreases the bias in inefficiency estimates by separating inefficiency from unobserved heterogeneity.[21] Therefore, the MREM is used here as the reference model and for comparison purposes, the PM and the REM are also estimated.[22]

Table 2 provides a summary of the model specification and a description of the stochastic terms included in the models. For each model, the estimate of each state's efficiency shown in the final row of Table 2 is based upon the conditional mean of the efficiency term proposed by Jondrow et al. (1982). The level of 'underlying energy

[20] Note that the Mundlak's formulation (i.e. with the introduction of this auxiliary equation in a REM) produces the 'Within Estimator'. In its original form, the Mundlak (1978) general panel data regression model is $Q_{it} = X_{it}\beta + AX_i\pi + \gamma_i + v_{it}$; however, Mundlak (1978) showed that the estimation of this model using GLS yields: $\hat{\beta}_{GLS} = \hat{\beta}_{within} \; and \; \hat{\pi}_{GLS} = \hat{\beta}_{Between} - \hat{\beta}_{within}$. The direct interpretation of the coefficients $\hat{\pi}_{GLS}$ is therefore not straightforward. Usually, the discussion on the results concentrate on $\hat{\beta}_{within}$.

[21] In this specification, it is assumed that the effect of unobserved state characteristics is captured by the coefficients of the group mean of the explanatory variables of equation (3).

[22] In a preliminary analysis, the TREM was also estimated. However, the results gave relatively high (and not really plausible) mean and median values of the efficiency level, i.e. 97% and 98%. As previously discussed the TREM produces efficiency estimates that do not include the persistent inefficiencies that might remain more or less constant over time and in this case, the results confirm the presence of this problem. For this reason, the TREM's results are not reported here.

efficiency', where the inefficiency varies over time (such as the PM), can be expressed in the following way:

$$EF_{it} = \frac{E_{it}^F}{E_{it}} = \exp(-\hat{u}_{it}) \tag{4}$$

where E_{it} is the observed energy consumption and E_{it}^F is the frontier or minimum demand of the i^{th} state in time t. An energy efficiency score of one indicates a state on the frontier (100% efficient); while non-frontier states, e.g. states characterized by a level of energy efficiency lower than 100%, receive scores below one. For the REM and the MREM the term u_{it}, in equation (4) is replaced by u_i and γ_i, respectively. This therefore gives the measures of 'underlying energy efficiency' estimated below.[23]

Table 2: Econometric Specifications of the Stochastic Energy Demand Frontier

	PM ***Half-Normal***	**REM** ***Half-Normal***	**MREM** ***Half-Normal***
State specific inefficiency u_i, u_{it} *and* γ_i	$u_{it} \sim$ N$^+$ (0, σ_u^2)	u_i N$^+$ (0, σ_u^2)	$u_i = AX_i\pi + \gamma_i$ $AX_i = \frac{1}{T}\sum_{t=1}^{T} X_{it}$ γ_i N$^+$ (0, σ_γ^2)
Random statistical noise v_{it}	$v_{it} \sim$ N (0, σ_v^2)	$v_{it} \sim$ N (0, σ_v^2)	$v_{it} \sim$ N (0, σ_v^2)
Efficiency	E($u_{it} \mid u_{it} + v_{it}$)	E($u_i \mid u_i + v_{it}$)	E($\gamma_i \mid \gamma_i + v_{it}$)

In summary, after estimating Equation (2), Equation (4) is used to estimate the efficiency scores for each state. The results from the estimation are given in the next section.

[23] This is in contrast to the alternative indicator of energy inefficiency given by the exponential of u_{it}. In this case, a value of 0.2 indicates a level of energy inefficiency of 20%.

4. Estimation results

The estimation results of the frontier energy demand models using the PM, the REM and the MREM are given in Table 3. The majority of the estimated coefficients and *lambda*[24] have the expected signs and almost all are statistically significant at the 10% level; the only exceptions being the share of detached houses in the REM and in the MREM and the coefficients of the auxiliary equation in the MREM.[25] Generally, the values of the estimated coefficients for the PM are different from those for the REM and MREM, whereas, the values of the estimated coefficients for some variables are different in the REM from the MREM. These differences are probably due to the problem of unobserved heterogeneity bias mentioned above. In fact, the values of the coefficients of the PM and the REM are different from those obtained by estimating equation (2) using a fixed effects estimator.[26] Whereas, the coefficients of the MREM are almost identical to the ones using the fixed effects estimator; hence, the most appropriate estimator for the efficiency analysis is considered to be the MREM.

Given that most of the variables are in logarithmic form, the coefficients can be directly interpreted as estimated elasticities. The results suggest that US residential energy demand is price-inelastic, with estimated elasticities of -0.07 -0.11 and -0.12 for the PM, the REM and the MREM respectively. The results also suggest that US residential energy demand is income-inelastic, with an estimated elasticity of 0.39 for the PM but only about 0.17 for the REM and 0.22 for the MREM. For weather, the estimated heating degree day

[24] Lambda (λ) gives information on the relative contribution of u_{it} and v_{it} on the decomposed error term ε_{it} and shows that in this case, the one-sided error component is relatively large.

[25] As already mentioned previously, the coefficients of the auxiliary equation in the Mindlak specification do not have a particular meaning. The goal of the variables included in the auxiliary equation is simply to reduce the unobserved heterogeneity bias. For a discussion of the Mundlak approach, see Baltagi (2006).

[26] In a preliminary analysis, equation (2) was estimated by using a classical fixed effects and random effects model. The results of the Hausman test confirmed the presence of the unobserved heterogeneity bias.

elasticities for all three models are about 0.4, whereas the estimated cooling degree day elasticities are rather low; ranging from 0.04 for the MREM to 0.08 for the PM. The estimated coefficient of average household size suggests that as family size increases, there is a tendency to use less energy; indicating there are economies of scale with an estimated elasticity of -1.11 for the PM, -0.55 for the REM, and -0.43 for the MREM. Whereas, for the share of detached houses, the results suggest that there is only a marginal positive influence on US residential energy demand; the estimated coefficient for the REM and the MREM being not significantly different from zero, and although for the PM the estimated coefficient is significantly different from zero, it is still rather low being 0.004.

Table 3: Estimated Coefficients (*t*-values in parentheses)

	PM	REM	MREM	
			Main equation	Auxiliary equation
Constant	-3.521 (-8.47)	-1.610** (-2.10)	-3.744 (-1.35)	
α^y	0.394*** (9.11)	0.166*** (3.44)	0.218*** (3.50)	0.172 (0.58)
α^p	-0.066** (-2.18)	-0.108*** (-3.71)	-0.118*** (-3.00)	0.097 (0.31)
α^{pop}	0.640*** (14.24)	0.855*** (16.53)	1.060*** (19.67)	-0.426 (-1.34)
α^{ahs}	-1.113*** (-15.94)	-0.554*** (-5.43)	-0.428*** (-5.07)	-0.737 (-1.28)
α^{hdd}	0.374*** (23.26)	0.420*** (16.43)	0.387*** (10.91)	0.026 (0.29)
α^{cdd}	0.088*** (10.72)	0.050*** (2.66)	0.035* (1.65)	0.046 (0.43)
α^{SDH}	0.004*** (8.14)	0.001 (0.20)	0.004 (1.08)	
Lamda (λ)	0.853*** (7.72)	5.686* (1.71)	4.368* (1.93)	
Sigma (σ)	0.103*** (822.6)	0.197*** (4.24)	0.147*** (4.17)	

***, ** and *: coefficients are significantly different from zero at the 99%, 95% and 90% confidence levels respectively.

For the PM and the MREM the time dummies, as a group, are significant and the overall trend in their coefficients is generally negative as shown in Figure 1. However, the estimated coefficients do not fall continually over the estimation period, reflecting the ‘non-linear’ impact of technical progress and other exogenous variables. The estimated coefficients for the REM have a ‘similar’ pattern to the PM and the MREM coefficients; nevertheless, a lot less individual coefficients are significantly different from zero.

Figure 1: Estimated Time Dummy Coefficients (relative to 19995)

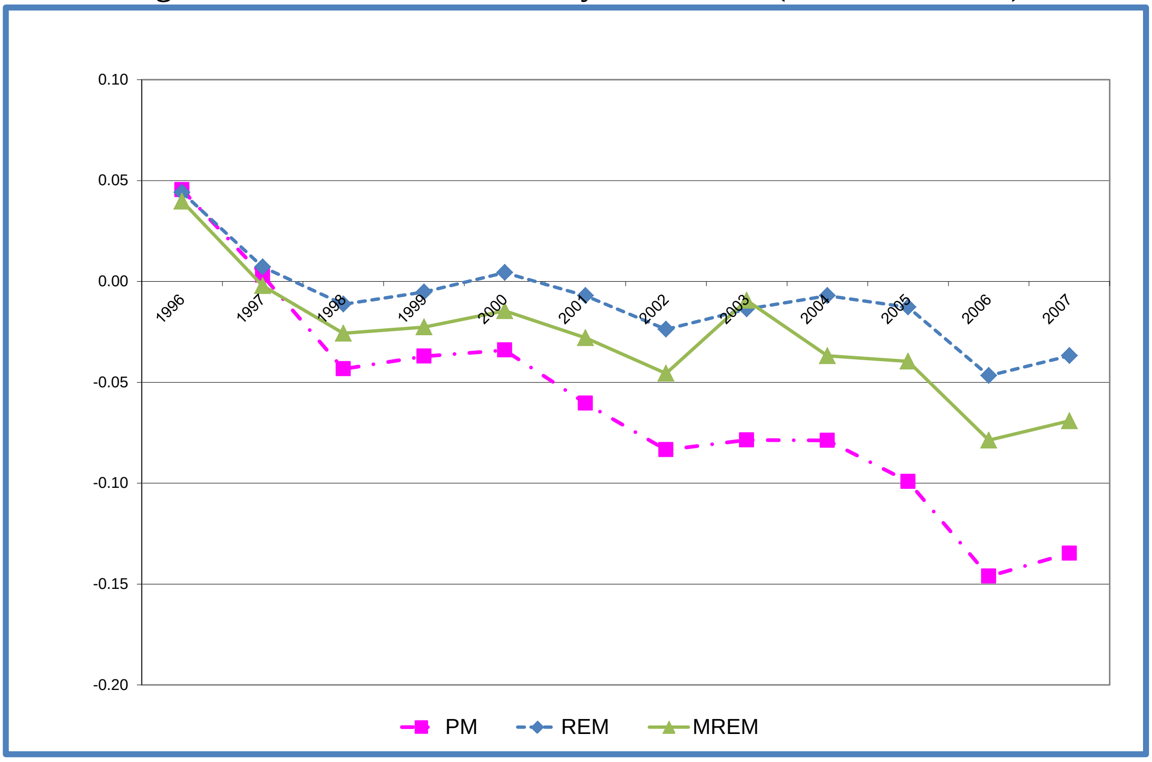


Table 4 provides descriptive statistics for the overall US estimated ‘underlying energy efficiency’ for the 48 states for the period 1995 to 2007. This shows that the estimated average efficiency is about 85% to 95%. As discussed above, given the length of the time series used in this research, the estimated coefficients from the PM and of the REM are likely to be affected by the so-called unobserved heterogeneity bias, and therefore the estimated levels of efficiency could be imprecise. Whereas, a shortcoming of the REM and the MREM results is that the estimated level of efficiency does not vary over time;

however, considering that the time period of the analysis is relatively short, this is acceptable. Furthermore, given that the heterogeneity bias in the MREM is minimal, all further analysis focuses on the results obtained using this model. Nevertheless, it is worth noting, that the Spearman's rank correlation coefficients between the three estimated levels of efficiency from the different estimates are 0.85 for the PM and the REM; 0.95 for the PM and the MREM; and 0.88 for the REM and the MREM. This indicates that the results in terms of the estimated level of efficiency tend to be robust across the different approaches and, although the MREM is taken as the preferred approach and further analysis is based on it, the results should also reflect those obtained from the PM and REM.

Table 4: Energy Efficiency Scores

	PM	REM	MREM
Min	0.87	0.64	0.72
Max	0.98	0.99	0.99
Mean	0.95	0.85	0.89
Median	0.95	0.85	0.90
st.dev.	0.02	0.08	0.07

As discussed in Filippini and Hunt (2011) it is expected that the estimated 'underlying energy efficiency' is negatively correlated with energy intensity; thus, when comparing across states, it is expected that the level of energy intensity decreases with an increase of the level of energy efficiency. However, as Filippini and Hunt (2011) argue, if this technique were to be a useful tool for teasing out 'underlying energy efficiency' then a perfect, or even near perfect, negative correlation would not be expected since all the useful information would be contained in standard energy intensity measures. This proves to be the case with the estimates here. The overall correlation coefficients between the estimated 'underlying energy efficiency' measure from the MREM and the average energy intensity measures suggested by the EIA (1995) report over the 1995 to 2007 period are -0.2 and -0.3 for 'energy per capita' and 'energy per building' respectively. Thus, as suggested, there appears to be a negative relationship, but it is by no means perfect. Nevertheless, of vital importance for US policy makers is the relative position across the states and if energy intensity were a good proxy for energy efficiency then there would need to be a high (positive) correlation between the rankings of the energy intensity measures and the estimated 'underlying energy efficiency' across the states. However, this is not the case with the Spearman's rank correlation coefficient across the 48 states for the period 1995 to 2007 being 0.2 for average 'energy per capita' and 0.3 for average 'energy per building'.[27] Table 5, Figure 2, and Figure 3 illustrate the rankings and clearly illustrate this relationship.

There are some states where the energy intensity measures would appear to be a good predictor of a state's rank of the estimated 'underlying energy efficiency', for both efficient and inefficient states. For example, Arizona is estimated to be the most efficient state according to the analysis above and is the state with the 3rd and 2nd lowest levels of average 'energy per capita' and average 'energy per building' respectively. Similarly, for Tennessee, which is estimated to be the 15th most efficient state according to the analysis above and the state with the 16th lowest levels of average 'energy per capita' and average 'energy per building' respectively. Also at the other end of the spectrum, Illinois is estimated to be the 47th most efficient state and is ranked 45th and 48th respectively according to the average 'energy per capita' and 'energy per building' measures.

However, there are also a number of states where the energy intensity measures would appear ***not*** to be a good predictor of a state's rank of the estimated 'underlying energy efficiency', for both efficient and inefficient states. For example, Florida is ranked 2nd and 1st respectively according to the average 'energy per capita' and 'energy per

[27] These two measures of energy intensity are used since the others suggested in the EIA (1995) report discussed in the introduction are not readily available.

building' measures, but is only 46th efficient according to the analysis above. Whereas Minnesota is ranked 40th and 41st respectively according to the average 'energy per capita' and 'energy per building' measures, but found to be relatively more efficient according to the analysis above, being ranked 3rd. Similarly, Louisiana is ranked the 7th and 9th most efficient state according to the average 'energy per capita' and 'energy per building' measures but is estimated to be only the 48th most efficient state according to analysis above.

Table 5: Comparison of the Rankings for Estimated Underlying Energy Efficiency (from the MREM) and Average Energy Intensity (1995-2007)

	Estimated Underlying Energy Efficiency		***Energy Intensity 1 (Energy per capita)***		***Energy Intensity 2 (Energy per building)***	
	Level	*Rank*	*Level*	*Rank*	*Level*	*Rank*
Alabama	0.849	33	36.722	14	83.311	11
Arizona	0.993	1	25.837	3	61.553	2
Arkansas	0.877	27	37.573	15	86.304	14
California	0.981	8	25.162	1	69.252	3
Colorado	0.992	2	42.020	24	98.377	25
Connecticut	0.790	43	51.577	44	126.304	46
Delaware	0.854	30	42.451	25	96.863	23
District of Columbia	0.918	19	40.415	21	84.933	13
Florida	0.767	46	25.656	2	55.649	1
Georgia	0.818	39	36.464	12	89.593	17
Idaho	0.970	9	38.273	18	93.386	19
Illinois	0.762	47	51.592	45	130.195	48
Indiana	0.854	30	47.477	37	112.154	40
Iowa	0.987	4	44.219	28	103.778	29
Kansas	0.869	29	46.344	33	108.385	35
Kentucky	0.877	27	40.470	22	93.596	20
Louisiana	0.719	48	33.935	7	81.555	9
Maine	0.827	37	58.892	48	115.466	42
Maryland	0.897	25	39.377	20	97.268	24
Massachusetts	0.815	41	48.428	41	116.997	45
Michigan	0.799	42	54.991	47	127.686	47
Minnesota	0.989	3	48.127	40	112.322	41
Mississippi	0.781	44	34.164	8	83.320	12
Missouri	0.909	22	45.147	30	102.750	28
Montana	0.906	23	45.202	31	101.168	26
Nebraska	0.913	21	47.252	36	110.711	38
Nevada	0.969	10	34.511	10	82.503	10
New Hampshire	0.927	18	47.488	38	106.235	31
New Jersey	0.818	39	45.714	32	115.739	43
New Mexico	0.985	5	33.667	6	78.951	6
New York	0.843	36	43.057	26	106.783	33

Table 5: Continued

North Carolina	0.984	6	35.203	11	79.732	7
North Dakota	0.933	16	50.347	43	108.948	36
Ohio	0.825	38	49.408	42	116.570	44
Oklahoma	0.847	34	41.712	23	94.828	22
Oregon	0.968	11	34.308	9	81.057	8
Pennsylvania	0.847	34	46.953	35	108.958	37
South Carolina	0.917	20	32.849	5	74.680	4
South Dakota	0.983	7	44.450	29	102.245	27
Tennessee	0.942	15	37.713	16	88.024	16
Texas	0.850	32	30.155	4	77.012	5
Utah	0.776	45	38.438	19	112.103	39
Vermont	0.899	24	51.661	46	105.923	30
Virginia	0.957	13	37.905	17	91.481	18
Washington	0.962	12	36.638	13	87.567	15
West Virginia	0.894	26	43.588	27	94.282	21
Wisconsin	0.955	14	46.870	34	107.004	34
Wyoming	0.930	17	47.828	39	106.293	32

Note: A rank of 48 for 'underlying energy efficiency' represents the least efficient state by this measure, whereas a rank of 1 represents the most efficient state. A rank of 48 for energy intensity represents the most energy intensity state whereas a rank of 1 represents the least energy intensive state.

Figure 2: Estimated Underlying Energy Efficiency (the MREM, 1995 - 2007)

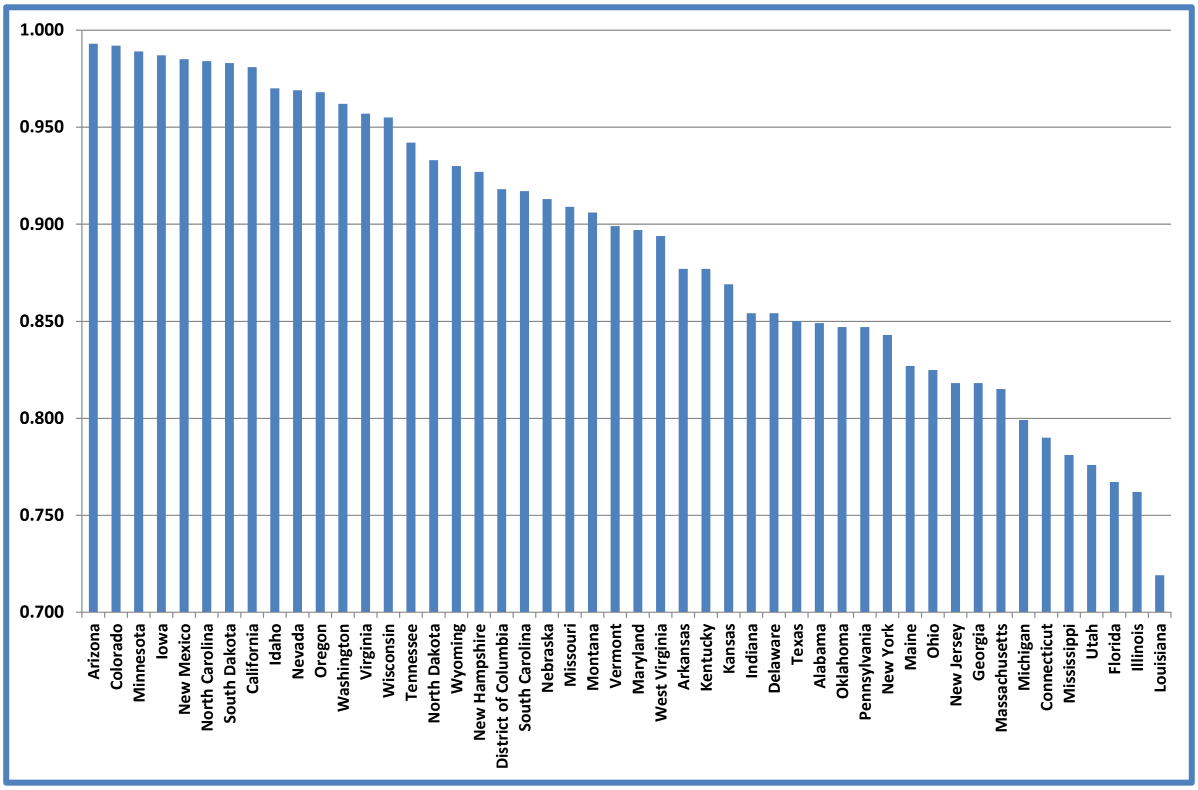
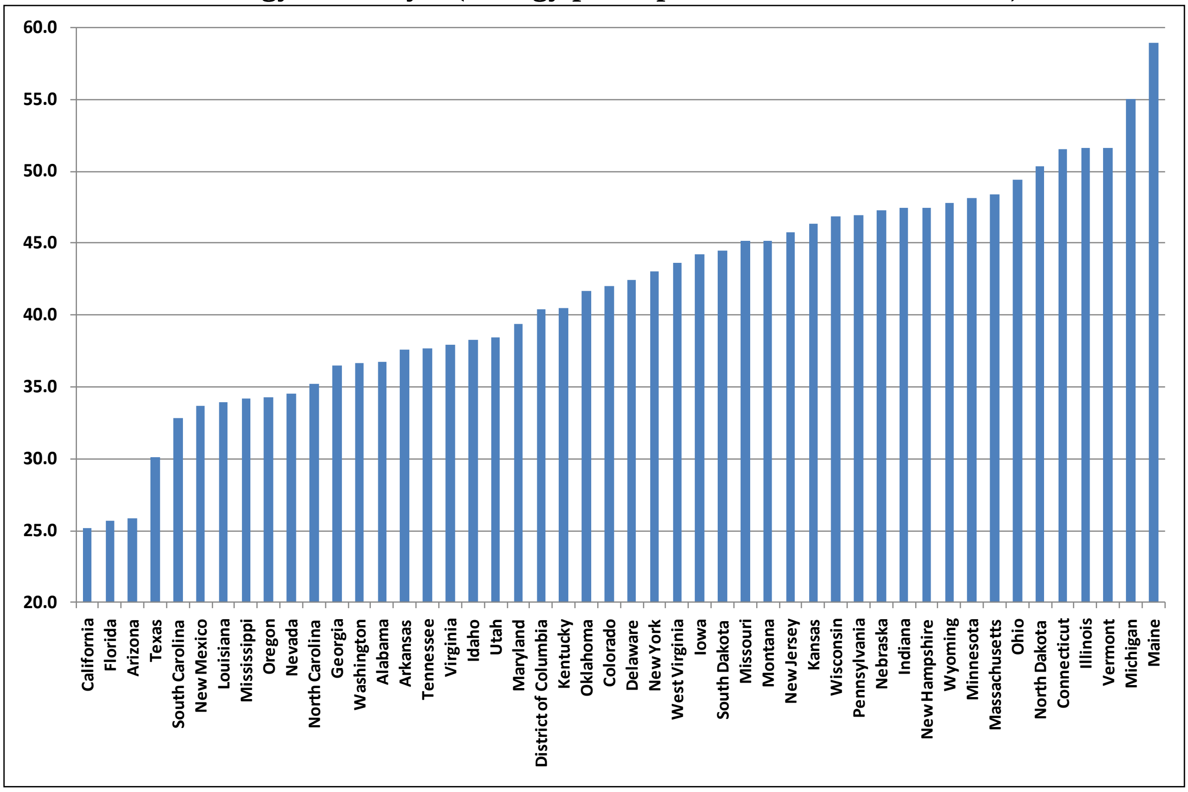
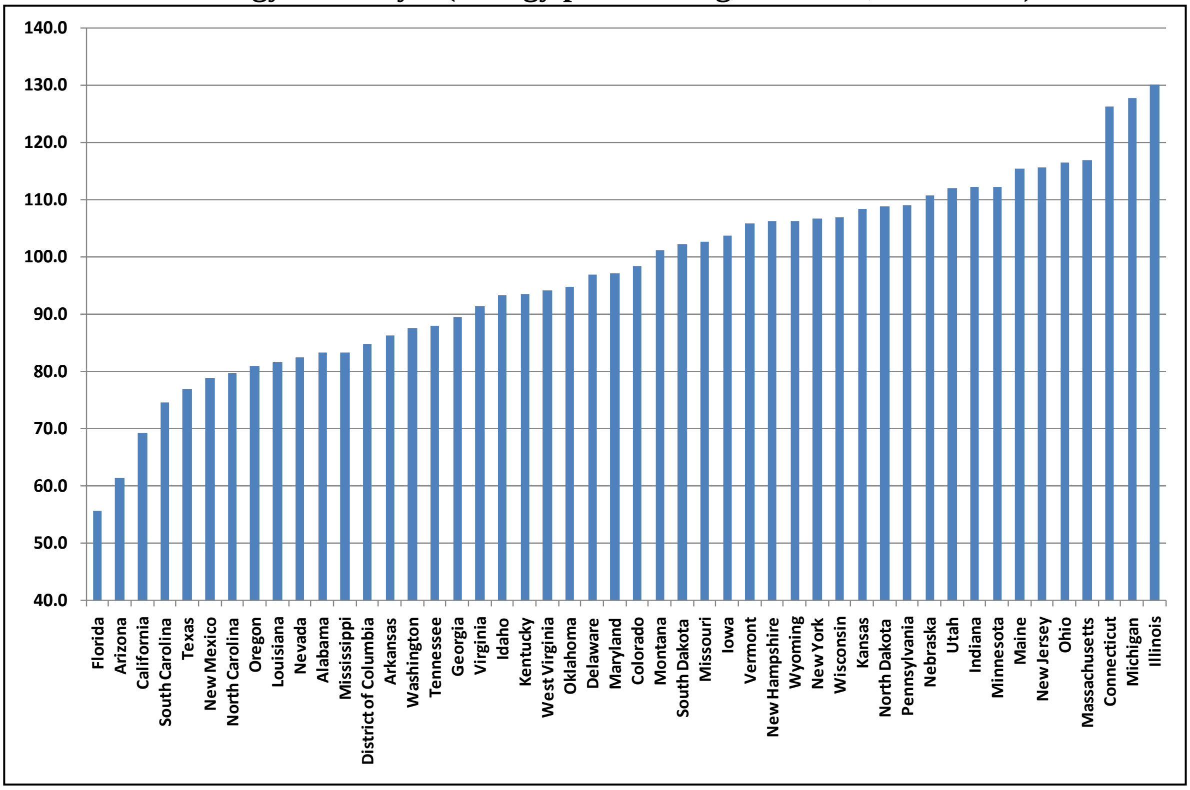


Figure 3: Average Energy Intensity

a: Energy Intensity 1 (Energy per capita, 1000 Btu, 1995-2007)



b: Energy Intensity 2 (Energy per building, 1000 Btu, 1995-2007)



5. Summary and Conclusion

Building on Filippini and Hunt (2011) this research attempts to isolate core US residential energy efficiency for a panel of 48 states, as opposed to relying on simple measures of energy intensity, such as 'energy per capita' or 'energy per building'. The approach taken combines energy demand modelling and frontier analysis in order to estimate the residential 'underlying energy efficiency' for each state. The energy demand specification controls for income, price, population, average household size, heating degree days, cooling degree days, the share of detached housing, and a UEDT in order to obtain a measure of 'efficiency' – in a similar way to previous work on cost and production estimation – thus giving a measure of residential 'underlying energy efficiency'.

The estimates for the underlying residential energy efficiency using this approach show that although for a number of states the change in the simple measures of energy intensity might give a reasonable indication of their relative energy efficiency (such as Arizona, Tennessee, and Illinois); this is not always the case (such as Florida, Minnesota and Louisiana). Therefore, unless the analysis advocated here is undertaken, US policy makers are likely to have a misleading picture of the real relative energy efficiency across the states and thus might make misguided decisions when allocating funds to various states in order to implement energy efficiency and conservation measures. Hence, it is argued that this analysis should be undertaken in order to give US policy makers an additional indicator to the rather naïve measure of energy intensity in order to try to avoid potentially misleading policy conclusions.

References

Adeyemi, O. I. and L. C. Hunt (2007) 'Modelling OECD industrial energy demand: asymmetric price responses and energy-saving technical change', ***Energy Economics***, 29, 693–709.

Aigner, D. J., C. A. K. Lovell and P. Schmidt (1977) 'Formulation and Estimation of Stochastic Frontier Production Function Models', ***Journal of Econometrics***, 6, 21-37

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

Baltagi B. H. (2006), 'An Alternative derivation of Mundlak's fixed effects results using system equation', ***Econometric Theory,*** 22, 1191-1194.

Banfi, S., M. Filippini, and Hunt, L. C. (2005),'Fuel tourism in border regions: the case of Switzerland', ***Energy Economics***, 27, 689–707.

Battese, G. E. and T. Coelli (1992) 'Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India', ***Journal of Productivity Analysis***, 3, 153–69.

Boyd G. A. and J. M. Roop (2004). 'A Note on the Fisher Ideal Index Decomposition for Structural Change in Energy Intensity', ***The Energy Journal,*** 25(1), 87-101.

Boyd G. A. (2008),'Estimating Plant Level Manufacturing Energy Efficiency with Stochastic Frontier Regression', ***The Energy Journal,*** 29(2), 23-44.

Buck, J. and D. Young (2007) 'The Potential for Energy Efficiency Gains in the Canadian Commercial Building Sector: A Stochastic Frontier Study,' ***Energy - The International Journal***, 32, 1769-1780.

Deaton, A. and Muellbauer, J. (1980), *Economics and consumer behavior.* Cambridge University Press, Cambridge, UK.

EIA (1995). *Measuring energy efficiency in the United States' economy: a beginning*. Energy Information Administration, DOE/EIA-0555(95)/2, Washington, DC, USA.

Farsi, M. and M. Filippini (2009) 'Efficiency Measurement in the Electricity and Gas Distribution Sectors', in J. Evans and L. C. Hunt (Eds) **International Handbook on the Economics of Energy**, Cheltenham, UK: Edward Elgar, pp. 598–623.

Farsi, M., M. Filippini and W. Greene (2005a) 'Efficiency Measurement in Network Industries: Application to the Swiss Railway Companies', ***Journal of Regulatory Economics***, 28, 69-90.

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

Farsi, M., M. Filippini and M. Kuenzle (2006) 'Cost Efficiency in Regional Bus Companies: An Application of Alternative Stochastic Frontier Models', ***Journal of Transport Economics and Policy***, 40, 95-118.

Filippini M., (1999) ‘Swiss residential demand for electricity’, ***Applied Economic Letters,*** 8, 533–538

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

Greene, W. (2005a) ‘Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier Model’, ***Journal of Econometrics***, 126, 269-303.

Greene, W. H. (2005b) ‘Fixed and random effects in stochastic frontier models’, ***Journal of Productivity Analysis***, 23, 7–32.

Griffin J. M. and C. T. Schulman (2005) ‘Price asymmetry in energy demand models: A proxy for energy-saving technical change?’, ***The Energy Journal***, 26(2), 1-21.

Hunt, L. C., G. Judge and Y. Ninomiya (2003a) ‘Underlying trends and seasonality in UK energy demand: a sectoral analysis’, ***Energy Economics***, 25, 93–118.

Hunt, L. C., G. Judge and Y. Ninomiya (2003b) ‘Modelling underlying energy demand trends’, Chapter 9 in: Hunt, L. C. (Ed.), ***Energy in a Competitive Market: Essays in Honour of Colin Robinson***, Cheltenham, UK: Edward Elgar, pp. 140–174.

Huntington, H G (1994) ‘Been top down so long it looks like bottom up to me’, ***Energy Policy***, 22, 833-838.

IEA (2009) ‘Progress with implementing energy efficiency policies in the G8’, ***International Energy Agency Paper***, www.iea.org/publications/free_new_Desc.asp?PUBS_ID=2127.

Jondrow, J., C. A. K. Lovell, I. S., Materov and P. Schmidt (1982) ‘On the Estimation of Technical Efficiency in the Stochastic Frontier Production Function Model’, ***Journal of Econometrics***, 19, 233-238.

Kumbhakar S.C. Lovell, C. A. (2000), *Stochastic frontier analysis*, Cambridge: Cambridge University Press.

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

Pitt, M. and L. Lee (1981) ‘The measurement and sources of technical inefficiency in the Indonesian weaving industry’, ***Journal of Development Economics***, 9, 43–64.

Schmidt, P. and R. E. Sickles (1984) Production frontiers and panel data’, ***Journal of Business and Economic Statistics***, 2, 367–74.

Zhou, P., B. W. Ang (2008) ‘Linear programming models for measuring economy-wide energy efficiency Performance’, ***Energy Policy,*** 36, 2911– 2916.

Zhou, P., B. W. Ang and K. L. Poh (2008) ‘A survey of data envelopment analysis in energy and environmental studies’, ***European Journal of Operational Research,*** 189, 1–18.