Measurement of Energy Efficiency Based on Economic Foundations

M. Filippini and L. C. Hunt

Economics Working Paper Series

ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Measurement of Energy Efficiency Based on Economic Foundations[#]

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

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

Surrey Energy Economics
Centre (SEEC),
School of Economics,
University of Surrey, UK

Abstract

Energy efficiency policy is seen as a very important activity by almost all policy makers. In practical energy policy analysis, the typical indicator used as a proxy for energy efficiency is energy intensity. However, this simple indicator is not necessarily an accurate measure given changes in energy intensity are a function of changes in several factors as well as 'true' energy efficiency; hence, it is difficult to make conclusions for energy policy based upon simple energy intensity measures. Related to this, some published academic papers over the last few years have attempted to use empirical methods to measure the efficient use of energy based on the economic theory of production. However, these studies do not generally provide a systematic discussion of the theoretical basis nor the possible parametric empirical approaches that are available for estimating the level of energy efficiency. The objective of this paper, therefore, is to sketch out and explain from an economic perspective the theoretical framework as well as the empirical methods for measuring the level of energy efficiency. Additionally, in the second part of the paper, some of the empirical studies that have attempted to measure the energy efficiency using such an economics approach are summarised and discussed.

JEL Classification: D, D2, Q, Q4, Q5.

Keywords: economic foundations of energy efficiency; energy demand; stochastic frontier analysis.

1   Introduction

This paper reviews the developments in attempting to define energy efficiency from an economics perspective and therefore how an economic based energy efficiency indicator should be measured/estimated.[1] There is considerable debate about the contribution of energy efficiency policies in enhancing energy security and helping reduce emissions from the use of energy. However, there is no consensus on how energy efficiency is actually defined and measured. The most typical indicator used in energy policy analysis is that of energy intensity, often defined as the simple ratio of energy consumption to GDP at the state or country level or energy consumption per square meter at the residential level.

According to EIA (1995; p.vii) without a clear definition, the term energy efficiency is "a vague, subjective concept that engenders directionless speculation and confusion rather than insightful analysis" and so there is a need to adequately define and measure what energy efficiency actually is, although recognising that this is a difficult task. Furthermore, EIA (1995) suggest that although energy intensity and energy efficiency are often used interchangeably, energy intensity does not necessarily reflect true energy efficiency given energy intensity is influenced by factors other than just pure energy efficiency. A view echoed by IEA (2009) who argue that using energy intensity as a proxy for energy efficiency is not appropriate given changes in energy intensity depend upon several factors such as the organization and structure of the economy and the real level of energy efficiency.

Over the past few years, some published academic papers have attempted to measure the level of the efficient use of energy based on the economic theory of production and used empirical methods for measuring productive efficiency. However, these studies have not generally provide a systematic discussion of the theoretical basis of energy efficiency. Neither have they clearly defined energy efficiency nor clearly shown how it should be empirically measured using parametric methods. The objective of this paper, therefore, is to sketch out and explain from an economic perspective the theoretical framework as well as the empirical methods for measuring the level of energy efficiency in order to try to deal with the problems highlighted by EIA (1995). Additionally, in the second part of the paper, some of the empirical studies that have attempted

---

[1] Note, that in previous works (Filippini and Hunt, 2011 and 2012) in order to highlight the distinction from using energy intensity as a measure of 'true' energy efficiency, we used the term 'underlying energy efficiency' to indicate the efficient use of energy obtained from estimating a frontier energy demand function using stochastic frontier analysis. We have not used this term here, but it should be remembered that when we refer to energy efficiency this relates to the estimated economic based energy efficiency indicator obtained from the techniques discussed in this paper and ***not*** measured (or proxied by) energy intensity.

to measure the energy efficiency using such an economics approach are summarised and discussed.

The paper is organized as follows. The next section presents and discusses productive efficiency and its relationship to energy efficiency. Section 3 then considers the developments in the parametric estimation of energy efficiency as defined here and the advantages and disadvantages of various stochastic frontier analysis models that can be applied. Moreover, a discussion of the distinction between permanent and transient productive efficiency is introduced. This is followed by Section 4 that highlights some of the attempts to measure energy efficiency, followed by a short summary and conclusion in the final section.

2. Productive efficiency and its relationship to energy efficiency[2]

There is no one clear and accepted definition of energy efficiency, but according to Bhattacharyya (2011), most definitions are based upon the simple ratio of 'useful output of a process/energy input into a process'. Furthermore, Patterson (1996) identifies a number of ways in which the outputs and inputs for this ratio can be quantified. These include: i) thermodynamic indicators (energy input and output are measured in thermodynamic units); ii) physical-thermodynamic indicators (energy input is measured in thermodynamic units and output is measured in physical units); iii) economic indicators (output and input are measured purely in terms of monetary values); and iv) economic-thermodynamic indicators (output measured in monetary values and the energy input measured in thermodynamic units). Moreover, these indicators can be applied at the product, sectoral or national levels of economic activity and for primary energy consumption and secondary final energy consumption.

Therefore, one of the most often used ratios in energy analysis at the macro level is the energy-GDP ratio, which is in fact the inverse of economic-thermodynamic indicator of energy efficiency identified by Patterson (1996). However, when undertaking aggregate energy efficiency analysis, this approach is arguably too simplistic and naïve and a better way to proceed is to use the definition based on the microeconomic theory of the production (see, Huntington 1994) advocated in Evans et al. (2013) and considered further below.

---

[2] This section builds upon and improves upon an initial attempt to consider these theoretical issues in Evans et al. (2013).

To understand this approach, it is necessary to recognise that the demand for energy is derived from the demand for outputs that can be products, services or energy services such as heating, cooling and lighting. Households and firms use energy, labour and capital to produce outputs. From an economic perspective, it is important to produce outputs in an efficient way; that is, by choosing the combination of inputs that minimise the production cost. In this context, situations where households or firms are producing outputs without minimising the use of inputs or using an obsolete technology that does not allow them to minimize the quantity of energy, labour and capital cannot be excluded. In these situations, the input energy as well the other inputs capital and labour are used in an inefficient way and a waste of energy is observed.

Productive inefficiency in the production of energy services can be discussed using the microeconomic theory of production framework with particularly isoquants and isocosts (Chambers, 1988; Huntington, 1994). In this context, the radial definition of technical, allocative and overall productive efficiency introduced by Farrell (1957) and particularly the non-radial concept of input specific technical efficiency introduced by Kopp (1981) can be helpful to understand the concept of energy efficiency

Figure 1 presents the situation of an economic agent that is using capital (C) and energy (E) to produce, in this case, an energy service (ES) such as heating.[3] The situation is illustrated using a unit isoquant (IS_0) and an isocost (IC_0) line. A technically efficient economic unit uses combinations of E and C that lie on the Isoquant IS_0.

---

[3] An economic unit could be a firm or a household and could refer to a wide range of energy services, such as heating, cooling, lighting, transportation, industrial processes, etc. Moreover, Figure 1 could also represent the economy wide aggregate production function for a state or a region.
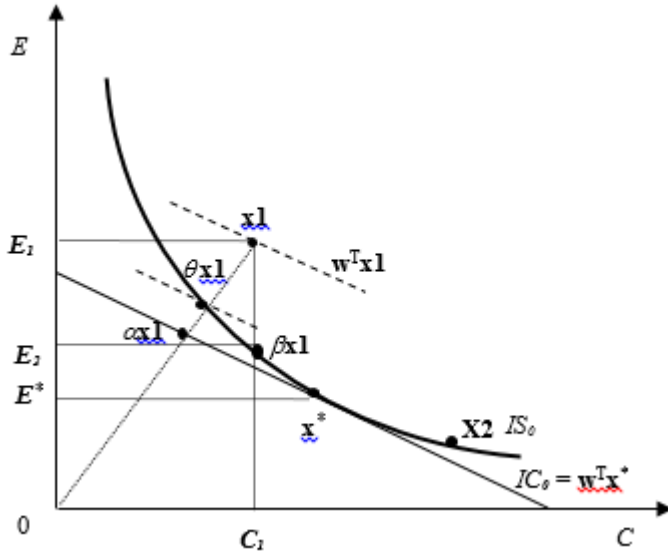
Figure 1: Productive efficiency

If the input price ratio, represented by the slope of the isocost line IC_0 in Figure 1 is known, then a cost efficient input combination can be identified. An economic agent that uses a cost-minimising input combination is illustrated by point $\mathbf{x}^*$, where the isocost line $\mathbf{w}^T\mathbf{x}^*$ is tangent to the isoquant IS_0, which reflects the production of a given level of energy services (ES*). Thus, the minimum costs required for the production of the given level of energy services (ES*) are $\mathbf{w}^T\mathbf{x}^*$.

If an economic agent uses a combination of inputs defined by point $\mathbf{x1}$ in Figure 1 to produce a level of energy service that correspond to the isoquant IS_0, it is technically inefficient as the point lies above IS_0. Using a classical input oriented radial measure, the level of technical efficiency θ can be measured as the ratio between the distance from the origin to technically efficient input combination $\theta\mathbf{x1}$ and the distance from the origin to input combination $\mathbf{x1}$. This measure treats the contribution of each input to technical efficiency equally (equi-proportionate).

From Figure 1 the economic agent operating at $\theta\mathbf{x1}$ is technically efficient but allocatively inefficient since it produces with higher costs. The level of allocative efficiency is measured as the ratio between the distance from the origin to $\alpha\mathbf{x1}$ and the distance from the origin to $\theta\mathbf{x1}$. The overall productive efficiency or cost efficiency α can be obtained as the ratio between the distance from the origin to $\alpha\mathbf{x1}$ and the distance from the origin to $\mathbf{x1}$. To reach the optimal input combination, the economic agent has to increase the use of input C and decreases the use of input

E. For example, an increase of *C* could be reached by installing a device on a cooling system to improve the function of the system or by substituting the single glazing windows with double glazing windows.

Note, that an economic agent that uses quantities of inputs defined by point **x1** it is technically as well allocative inefficient. The economic agent could improve the level of overall productive efficiency by moving to the optimal input combination **x***. In this case, energy consumption will decrease, as energy is substituted with capital and used in a more parsimonious way allowing the economic agent to consume less energy. This occurs for instance when a household or a firm improve the insulation of the building, change some electrical appliances, optimize the use of the heating or cooling system and the use of electrical appliances, in order to reach **x***.

So far, a radial notion of technical efficiency has been discussed. In this case, an improvement of the level of efficiency in the use of inputs requires a reduction in energy and the other inputs proportionally. However, if researchers are more interested in obtaining an input specific technical efficiency measure, for example, as here, in terms of an energy efficiency measure, then empirical analysis should be based on a non-radial notion of efficiency. This measure, introduced by Kopp (1981), can be expressed as the ratio between the distance from the technically efficient input vector β**x1** and the input vector **x1** in Figure 1. This is a special case of technical efficiency and, considering the input energy, is defined as the ratio of minimum feasible (E_2) to observed use of energy (E_1), conditional on the production technology and the observed levels of outputs and other inputs.

It is worth noting that, as discussed in Kopp (1981), all the theoretical definitions presented in Figure 1 help to understand the concept of technical, allocative and input specific efficiency. However, empirical studies on productive efficiency provide measures of efficiency that represent a generalization of these theoretical definitions. For example, the assumption of a unit isoquant is usually relaxed in empirical studies. Furthermore, with empirical frontier production, cost or input demand functions the level of efficiency is measured as deviations in output, input or cost from these frontiers.

As discussed in more detail in the next section, Filippini and Hunt (2011), motivated by the notion of non-radial input specific efficiency introduced by Kopp (1981), proposed that the way to econometrically estimate a measure of the efficient use of energy is to estimate a single

conditional input demand frontier function, such as a demand function for energy. With this approach the difference between the optimal use of energy ($E*$) that corresponds to the cost minimizing input combination x* to produce a given level of energy services (ES*) and the observed use of energy (E_1) should be measured.

In summary the above suggests that the non-radial measure of input specific efficiency is an interesting concept to use in the empirical analysis of the level of the efficient the use of energy for the production of outputs or energy services. As will be seen later in the paper, the majority of the empirical studies dealing with the measurement of energy efficiency are based on this non-radial measure of input specific efficiency.

In the analysis of the productive efficiency of an economic agent it is also important to consider the impact of technological change on the efficient use of energy and, more generally, on the use of all inputs. In this context, when technological change allows the economic agent to produce the same level of the energy service ES*, with less energy and capital, the isoquant shifts. For instance, this occurs when the temperature of rooms in a house is maintained at say 20° Celsius, with less energy and capital maybe due to new insulation technology or a new heating system. In this case, the technological progress will move the isoquant, IS$_0$, to the left as depicted in Figure 2.[4] In this case, the amount of energy and capital used to produce the energy service has decreased and the economic agent reaches point **xt1***.[5] If an economic agent decides to use an obsolete technology and remains at **xt0*,** then a technological gap is observed, i.e. the economic agent does not adopt a new technology that allows the inputs to be reduced and therefore decrease cost. [6]

---

[4] See also Gillingham et al. (2009) for a discussion of this effect.

[5] In Figure 2, a homothetic production function is assumed, which implies a parallel shift of the isoquant.

[6] The use of an obsolete technology is related to the 'energy efficiency gap' concept often discussed in the energy economics literature (see, for example, Jaffe and Stavins, 1994). The energy efficiency gap is a situation where households or firms do not perform an investment in a new energy saving technology, although from an economic point of view the investment is profitable and sustainable. We think that from a microeconomic point of view the term technology gap is more appropriate, because it is broader and refers to a new technology
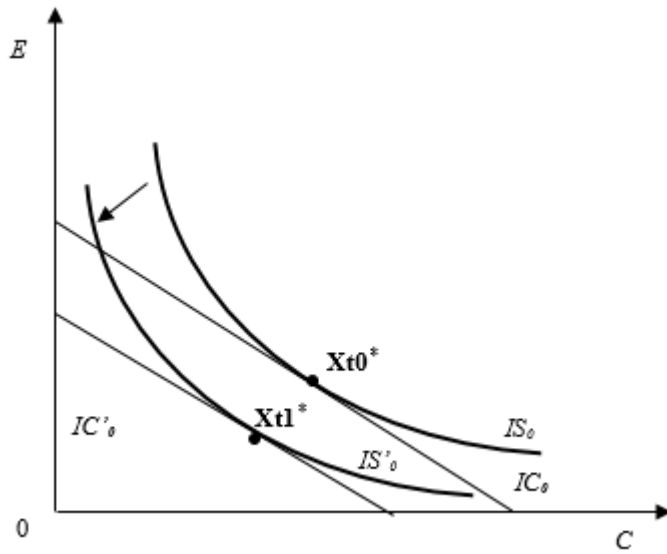
Figure 2: Technical progress

From this discussion, it is clear, that the level of energy consumption for the production of a predefined level of energy service can change over time because of an increase of the level of the productive efficiency, input specific efficiency and/or technical progress. However, as Kumbhakar and Lovell (2000) note, when using parametric methods to estimate production frontier functions the distinction between the impact of the increase in the level of productive efficiency and the impact of the introduction of new technologies on output can be difficult.

Finally, it is worth noting that the potential explanations for the presence on inefficiency in the use of energy or the presence of a technological gap tend to fall into two categories: market failures such as information problems, misplaced incentives and behavioural failures such as heuristic decision making, inattentiveness and bounded rationality.[7]

This section has focussed on the theoretical underpinnings of defining energy efficiency from an economics perspective. The following section therefore builds on this by considering the range of empirical parametric methods that could be considered for estimating energy efficiency based on this definition.

---

[7] See Gillingham et al. (2009) for an interesting presentation of the different type of market and behavioural failures.

3. Parametric methods to energy efficiency estimation

In order to estimate the level of overall productive efficiency, the level of efficiency in the use of a single input and the level of technical progress it is possible to use parametric and non-parametric frontier analysis.[8] In non-parametric approaches the production or cost frontier function is considered as a deterministic function of some variables but no specific functional form is imposed. Moreover, non-parametric approaches can be applied on small datasets. On the other hand, parametric approaches based on econometric methods are able to take into account, at least partially, unobserved heterogeneity among different economic agents but need to specify a functional form for the cost, production or input demand function. The main advantage of parametric approaches over non-parametric approaches is the separation of the inefficiency component from the statistical noise due to data errors and omitted variables, and accommodates formal statistical testing. Moreover, using panel data and some specific stochastic frontier models it is possible to reduce the unobserved heterogeneity bias because for the estimation of empirical frontier models some explanatory variables are usually not observed. The non-parametric methods' assumption of a unique deterministic frontier for all production units is strong. In this paper, we suggest that parametric methods are more attractive to analyse the level of efficiency in the use of energy, because of the presence in the modelling of unobserved heterogeneity in the production of energy services; nevertheless, we do recognize the possible advantages of using the non-parametric approach.[9] However, in the following sections the use of parametric methods for the estimation of frontier functions only are discussed and, given the majority of the parametric methods have a stochastic element in their frontier function, this group of methods are generally known as Stochastic Frontier Analysis (SFA).[10]

For the estimation of an input specific technical efficiency measure, e.g. a measure of the efficient use of energy, three approaches can be used. Reinhard et al. (1999) show how to obtain, through a two-step procedure, an indicator of input specific technical efficiency from the estimation of a production frontier. Kumbhakaker and Hjalmarrson (1995), on the other hand, propose the measurement of input specific technical efficiency from the estimation of an input

---

[8] See for instance, Murillo-Zamorano (2004) for a general presentation of different methodologies.

[9] In the literature it is possible to find several applications of non-parametric approaches such as Data Envelopment Analyis (DEA) that attempt to measure the efficient us of energy; see for example, Hu and Wang (2006), Zhou and Ang (2008) and Zhou et al, (2012).

[10] See Kumbhakar and Lovell (2000) and Greene (2008) for a discussion on the estimation of production, cost and input distance functions.

requirement function. This function indicates the minimum amount of an input that is necessary to produce a given level of output, given the technology and the quantity of other production factors. Finally, Zhou et al. (2012) propose the estimation of a Shephard energy input distance function, and Boyd (2008) has shown that this type of input distance function is equivalent to the input requirement function.[11] Furthermore, it is worth highlighting that all three approaches are able to estimate only the level of technical efficiency in the use of an input; however, from an economic point of view it is important to also have information on the level of overall efficiency in the use of an input, i.e. a measure that includes both technical and allocative components of inefficiency.

Filippini and Hunt (2011) suggested that the estimation of a measure of the efficient use of energy could also be based on the estimation of a single conditional input demand frontier function, such as the demand function for energy.[12] This function indicates the minimum amount of energy that is necessary to produce a given level of output (energy services), given the technology, input prices and other factors. In this context, actual energy demand differs from the energy frontier demand due to the presence of both allocative and technical inefficiency and not just technical inefficiency, as in Kopp (1981).[13] Nonetheless, it should be noted that the approach proposed by Filippini and Hunt (2011) should be considered an ad-hoc approach because it does not completely consider the theoretical restriction imposed by production theory and just estimates one input demand frontier function.[14] Therefore, this approach can only estimate an approximation of the inefficiency in the use of energy, determined by the presence of both allocative and technical inefficiency. However, in principle, this approach provides information on the difference between the optimal use of energy that corresponds to a cost minimizing input combination to produce a given level of energy services and the observed use of energy. This indicator of the efficient use of energy, measures the ability of an economic agent to minimize

---

[11] Sometimes, the Shephard energy distance function is also called sub-vector energy distance function

[12] Schmidt and Lovell (1979) use the self-duality of the Cobb-Douglas production function to derive a system of log-log stochastic cost-minimizing input demand equations, which are also known as input frontier equations.

[13] See Kumbhakar and Lovell (2000) for a discussion on this point.

[14] Schmidt and Lovell (1979) propose to estimate the cost frontier function together with all input demand frontier functions. This approach satisfies the theoretical restriction imposed by production theory and take into account simultaneously the fact that the input allocative efficiency can be different in each input demand frontier function. Of course, also the other approaches to measure energy efficient are based on the estimation of just one equation.

the energy consumption to produce a level of energy services given input prices and one of the advantages of this approach is that it considers the cost-minimizing input combination.

However, as discussed by Schmidt and Lovell (1979), the sign of the allocative efficiency in an input demand function can be positive or negative, i.e. the under or over use of an input can be observed. In order to reach an allocative efficient input combination the economic agent might have to change the proportions of inputs, given their relative prices. This change of the proportions could also determine an increase of the consumption of some inputs and not always a reduction. This implies that by using the approach introduced by Filippini and Hunt (2011) based on the estimation of just one input demand function, the optimal use of energy (E^*) is not always necessarily lower than the observed use of energy (E_1). For instance, point **x2** in Figure 1 is characterized by an inefficient input combination. In this case, in order to reach **x*** it is necessary to substitute capital with energy, i.e. energy consumption will increase. This means that in an empirical study, inefficient input combinations could be identified where energy is overused and combinations where energy is underused, i.e. the sign of the part of allocative efficiency can be positive or negative. Which situation dominates in order to reach a cost-minimizing input combination (energy reducing and other inputs increasing or energy increasing and other inputs reducing) is purely an empirical question. Of course, for the estimation of an input demand function using SFA, which is based on the presence of a strictly nonnegative component (a measurement of inefficiency), the presence of over use or under use of energy is likely to result in econometric problems, such as wrong skewness of the nonnegative component. That said, the fact that economic agents using obsolete technologies/appliances can be observed, tends to support the hypothesis that the optimal use of energy (E^*) is generally lower than the observed use of energy (E_1). Nonetheless, the potential limits of this approach should be always considered in any empirical analysis and in the interpretation of the results.

In summary, it is believed that all four approaches just illustrated are a sounder basis for measuring energy efficiency based on economic foundations rather than relying on SIMPLE energy intensity indicators. However, as discussed in more detail later, the actual empirical studies that estimate the level of efficiency in the use of energy are generally based on the estimation of three functions: i) an input requirement function (Boyd, 2008); ii) a Shephard energy distance function (Zhou et al., 2012); and iii) an energy demand function (Filippini and Hunt, 2011). Therefore, the following discussion is related only to these three functions and not on the approach proposed by Reinhard et al. (1999).

From an econometric perspective, the estimation of these frontier functions is based on variants of the SFA frontier function approach proposed by Aigner et al. (1977). This approach is based on the assumption that the level of energy inefficiency can be approximated by a one-sided non-negative term. Below, therefore provides details of the econometric specifications of the three functions that can be used to estimate the level of energy efficiency.

3.1 Input Requirement Function Econometric Specification

Using a log-log functional form, considering panel data and adopting the panel version of the original SFA frontier function approach proposed by Aigner et al. (1977) an input requirement frontier function for energy can be specified as follows:

$$ln\text{E}_{it} = \alpha + \boldsymbol{\beta}'\mathbf{x}_{it} + \varepsilon_{it}$$
$$v_{it} \sim N[0, \sigma_v^2],$$
$$u_{it} = |U_i|, \ U_{it} \sim N[0, \sigma_u^2], \tag{1}$$
$$\varepsilon_{it} = v_{it} + u_{it},$$

where i denotes the economic unit and t the year. E_{it} is energy consumption, x_{it} is a set of time-varying and time-invariant covariates such as the output and the inputs in logs. Further, x_{it} can also include a fully set of time dummies or a time trend to capture technological change.[15] β is the associated vector of parameters to be estimated. The error term ε_{it} in Equation (1) is split into two independent parts. The first part, u_{it}, reflects the energy efficiency and is interpreted as an indicator of the level of inefficiency in the use of energy. It is a one-sided non-negative random disturbance term that can vary over time. Moreover, in the empirical applications it is generally assumed that u_{it} follow a half-normal distribution. The second part, v_{it}, is the classical

---

[15] In order to consider the effect of technological change on production, cost or input demands, all frontier models can be estimated by introducing a set of time dummy variables. Generally, it is assumed that these variables capture the shift in the frontier functions due to change in the technology. Of course, as suggested by the energy demand literature, such shifts might also come about by changes in a range of exogenous impacts that cannot usually be measured explicitly, consistent with the idea of an Underlying Energy Demand Trend (or UEDT for short). The impact of such technological, organizational, and social innovation in the production and consumption of energy services on the energy demand can be also captured with a time trend, but this a more restrictive version of the UEDT (Hunt et al., 2003a & 2003b).

symmetric disturbance capturing the effect of noise and as usual is assumed to be normally distributed.[16]

3.2 Shephard Input Distance Function Econometric Specification

In the case where the interest is in the estimation of a Shephard input distance frontier function, then the following model can be specified:

$$
\begin{aligned}
&ln(\frac{1}{E_{it}}) = \alpha + \boldsymbol{\beta}'\mathbf{x}_{it} + \varepsilon_{it} \\
&v_{it} \sim N[0, \sigma_v^2], \\
&u_{it} = |U_i|, \ U_{it} \sim N[0, \sigma_u^2], \\
&\varepsilon_{it} = v_{it} - u_{it},
\end{aligned}
\tag{2}
$$

The differences between the input requirement functions and the Shephard input distance function are the dependent variables and the sign of the inefficiency term.

3.3 Input Demand Frontier Function Econometric Specification

Finally, researchers interested in the estimation of an input demand frontier function can use the following specification:

$$
\begin{aligned}
&ln E_{it} = \alpha + \boldsymbol{\beta}'\mathbf{z}_{it} + \varepsilon_{it} \\
&v_{it} \sim N[0, \sigma_v^2], \\
&u_{it} = |U_i|, \ U_{it} \sim N[0, \sigma_u^2], \\
&\varepsilon_{it} = v_{it} + u_{it},
\end{aligned}
\tag{3}
$$

where in this case z_{it} is a set of time-varying and time-invariant covariates such output,inputs prices and time dummies. As in the previous model, u_{it} reflects the level of energy efficiency. The difference between the input requirement function and the input energy demand function is that the former includes the quantity of inputs whereas the latter the input prices.

---

[16] This assumption allows the 'identification' of the efficiency for each economic unit separately. This is a standard assumption used in the estimation of production or cost frontier functions; see Kumbhakar and Lovell (2000, p. 148) for a discussion.

3.4 Econometric Specification and Estimation

A common econometric specifications for equations (1), (2) and (3) is the log-log functional form, although there are other possibilities with a high number of alternative functional forms proposed in the literature.[17] Generally, a distinction can be made between traditional and flexible functional forms; the most important difference being that the former imposes restrictions on the values of the first and second partial derivatives whereas the latter do not. One consequence of this being that for the log-log functional form used in equations (1), (2) and (3), the price elasticity is constant, not varying with demand; whereas flexible functional forms, such as the translog, allows for the price elasticity to vary with demand.[18] Still, it is worth noting that while the flexible functional forms such as the translog has many desirable attributes, the applied economists should always keep in mind the objective of their research when choosing a functional form; relevant considerations being the number of explanatory variables to be considered in the model, the number of observations and the multicollinearity problems that may arise using a flexible functional form.

Equations 1, 2 and 3 are specified for panel data but the estimation of such a stochastic frontier function can also be performed using cross-sectional data. However, panel data sets have the advantage that they allow the use of econometric models that can take into account the possible presence of unobserved heterogeneity in the model specification. Another econometric problem that should be considered, especially in the estimation of an input requirement function and a Shephard energy distance function, is the potential endogeneity problem of inputs and/or outputs.[19]

---

[17] For a good overview of functional forms in applied production analysis, see Chambers (1988).

[18] Within a group of flexible functional forms it is possible to differentiate between those functional forms derived from second-order Taylor series approximations, such as the translog, the generalized Leontief and the quadratic or the functional forms derived from Fourier or Laurent series approximations. The translog form provides a second-order approximation of the true function at a given point. Therefore, the independent variables are expressed as deviations from this point. Often, the sample mean or median of each of the explanatory variables is chosen.

[19] For a discussion on this issue, see Guan et al. (2009).

The econometric models available for estimating a stochastic frontier model using panel data are numerous.[20] In this paper, discussion on some of these models is restricted to those most commonly used and applicable to the task of estimating energy efficiency. These are: the pooled model (PM hereafter) that is basically the panel version of the model proposed by Aigner et al. (1977) and presented above;[21] the random effects model (REM hereafter); the true random effects model (TREM hereafter);[22] and the 'Generalized True Random Effects' (GTREM hereafter). As discussed in Filippini and Greene (2015), some of these models produce time invariant values of the level of efficiency (persistent efficiency), whereas others models estimate time variant values (transient efficiency). We believe that this distinction in the type of efficiency is relevant and will gain importance in future research. Table 1 summarizes the econometric specifications shortly discussed in this paper.

---

[20] For an overview of all the models for panel data, see Kumbhakar et al. (2012) and Filippini and Greene (2015).

[21] Battese and Coelli (1995) proposed a basic pooled model where the inefficiency term u_{it} is modified so as to have a systematic component associated with a vector of variables ($\mathbf{z}_{it}$) and a random component (e_{it}): $u_{it} = \eta' \mathbf{z}_{it} + v_{it}$. using panel data. However, it should be noted that this model does not in fact exploit the panel aspect of the data set in order to deal with the unobserved heterogeneity.

[22] See Greene (2008) and Farsi and Filippini (2009) for a discussion of these models.

Table 1: Econometric Specifications of the Stochastic Cost Frontier

	Model I **PM** (Pitt and Lee)	*Model II* **REM** (Pitt and Lee)	*Model III* **TRE**	*Model IV* **GTREM**
Firm effects α_i	None	α	$N(\alpha,\sigma_w^2)$	$N(\alpha,\sigma_w^2)$
Full random error ε_{it}	$\varepsilon_{it}=u_{it}+v_{it}$ $u_{it}\sim N^+(0,\sigma_u^2)$ $v_{it}\sim N(0,\sigma_v^2)$	$\varepsilon_{it}=u_i+v_{it}$ $u_i\sim N^+(0,\sigma_u^2)$ $v_{it}\sim N(0,\sigma_v^2)$	$\varepsilon_{it}=w_i+u_{it}+v_{it}$ $u_{it}\sim N^+(0,\sigma_u^2)$ $v_{it}\sim N(0,\sigma_v^2)$ $w_{it}\sim N(0,\sigma_w^2)$	$\varepsilon_{it}=w_i+h_i+u_{it}+v_{it}$ $u_{it}\sim N^+(0,\sigma_u^2)$ $h_i\sim N^+(0,\sigma_h^2)$ $v_{it}\sim N(0,\sigma_v^2)$ $w_i\sim N(0,\sigma_w^2)$
Persistent Inefficiency Estimator	None	$E(u_i\,|\,\varepsilon_{i1},\ldots\varepsilon_{iT})$	None	$E(h_i\,|\,\mathbf{y}_i)$
Transient Inefficiency Estimator	$E(u_{it}|\varepsilon_{it})$	None	$E(u_{it}|\varepsilon_{it})$	$E(u_{it}\,|\,\mathbf{y}_i)$

The PM does not exploit the possibility given by panel data to control for time invariant unobserved heterogeneity variables. Therefore, the unobserved heterogeneity bias can be present in this model. On the other hand, the REM introduced by Pitt and Lee (1981) considers the typical panel data individual random effects as inefficiency rather than unobserved heterogeneity as in the literature on panel data econometric methods.[23] The REM produces efficiency indicators that do not vary over time (persistent efficiency). One problem with the REM is that unobserved time invariant variables are captured by the individual effects and, therefore, considered as inefficiency. As an alternative to the REM, Greene (2005a and 2005b) proposed the TREM

---

[23] Several variants of the Pitt and Lee (1981) model have been proposed to accommodate time variation in the inefficiency term. Generally, in these models, the inefficiency term is represented as a product of a deterministic function of time and the random effects, u_i. For instance, Kumbhakar (1990) specifies the inefficiency term as $u_{it}=[1+\exp(bt+ct^2)]^{-1}\,|\,U_i\,|$, Battese and Coelli (1992) as $u_{it}=\exp[-\eta(t-T)]\,|\,U_i\,|$, Battese and Coelli (1995) as $u_{it}=\exp[g(t,T,\mathbf{z}_{it})]\,|\,U_i\,|$.

whereby the PM is extended by adding fixed and random individual effects respectively.[24] The TREM produce values of the level of efficiency that vary over time (transient efficiency). In these models the individual fixed or random effects take into account all unobserved variables that are time-invariant. In the TREM any time-invariant or persistent component of inefficiency is completely absorbed in the state-specific constant terms. Therefore, if the production of energy services is characterized by persistent inefficient use of energy (for instance by the presence of old production machines, old buildings or an old road system or systematic behavioural failures such as adjusting the room temperature or frequently opening the windows), the TREM will produce relatively high levels of estimated energy efficiency. These approaches have been further developed more recently with some scholars proposing the GTREM that allows for the possibility of estimating the level of persistent and transient efficiency of an economic agent at the same time. The GTREM recognizes that the level of productive efficiency can be split into two components and the importance of estimating both components at the same time.[25] This is obtained by adding, to the TRE model, a time persistent counterpart to u_{it} in the time varying stochastic frontier. Therefore, this model includes a four-part disturbance with two time varying components and two time invariant components as presented in Table 1.

Finally, it is worth noting that all the econometric approaches outlined above can suffer from 'unobserved variables bias', because the unobserved characteristics may not be distributed independently of the explanatory variables. In order to address this econometric problem Farsi et al. (2005b) proposed that a version of these models should be estimated using the procedure proposed by Mundlak (1978), whereby the correlation of the individual specific effects and the explanatory variables are considered in an auxiliary equation.[26]

Given, this discussion the following section considers some of the published work that has attempted to apply some of these techniques for estimating the efficient use of energy.

---

[24] Greene (2005) has also proposed the True Fixed Effects Model. A model similar to the TREM with the difference that the individual effects are fixed and not random.

[25] For instance, Kumbhakar et al. (2012) and Colombi et al. (2014) have proposed a relatively complex approach that provide separate estimates of the two components of efficiency. More recently, Filippini and Greene (2015) have proposed a relatively straightforward estimation method.

[26] See Filippini and Hunt (2012) for an application of this approach.

4   Empirical applications and extensions

The literature on the econometric estimation of the efficient use of energy is relatively new, but several studies have now been published based on the approaches outlined in the previous sections – some using aggregate date and a few using disaggregate data.[27] The previous section detailed that there are effectively three approaches to econometrically estimating the level of efficiency in the use of energy. The first approach involves the estimation of an energy requirement function[28] (that provides information only on technical efficiency), the second involves the estimation of a Shephard energy distance function[29] (that provides information only on technical efficiency) and the third involves the estimation of an energy demand frontier function[30] (that provides information on both technical and allocative efficiency). A selection of papers using all three approaches are reviewed below.[31] As is shown, the empirical studies discussed in this section differ in several dimensions; such as in the type of function used, in the type of data (cross sectional or panel) employed and in the type of efficiency (persistent and transient) estimated. In fact, the majority of these studies provide a measure of energy efficiency that does not distinguish between transient and persistent energy efficiency given that, as discussed previously, only recently has this differentiation been introduced into the literature on the measurement of productive efficiency.

4.1 Examples of Energy Requirement Function Estimation

Boyd (2008) is an early example of research estimating an energy input requirement using SFA to analyse the efficient use of energy in the industrial sector. Boyd (2008) estimates a cross section model with 37 observations (plant years) based on data spanning 1992 to 1997 for a non-public micro-dataset of wet corn milling plants. Boyd's (2008) specifies primary energy use as being dependent upon variables representing the amount of corn processed, the amount of

---

[27] There are also examples of studies that estimated the level of efficiency in the use of other natural resources such as water and nitrogen, such as Reinhard et al. (1999) and Karagiannis et al. (2003), but they are not considered here given the focus on measuring energy efficiency.

[28] Note that in the previous section this was referred to an input requirement function, but in this section as an energy requirement function given the focus of this review.

[29] Note that in the previous section this was referred to a Shephard input distance function, but in this section as a Shephard energy distance function given the focus of this review.

[30] Note that in the previous section this was referred to an input demand frontier function, but in this section as an energy demand frontier function given the focus of this review.

[31] Note, the notation used in this section is wherever possible as close to the notation used in the papers that are reviewed.

modified starch produced, the amount of anhydrous dextrose, the amount of glucose syrup sweeteners and solids, the total amount of alcohol, the moisture content of gluten feed, and capacity utilization. In addition, a white noise error term and a non-negative error term capturing the inefficient use of energy are included in the equation and assumed to be truncated normal. From this Boyd (2008; p. 40) claims that this approach "is an important contribution to the measurement of energy efficiency, particularly in manufacturing where plants may produce products that vary in their energy requirements."

Another example of using an energy requirement function is Lin and Wang (2014) who estimate a translog specification as follows:[32]

$$ln\, E_{it} = \beta_0 + \beta_y lnY_{it} + \beta_k lnK_{it} + \beta_l lnL_{it} + \frac{1}{2}\beta_{yy}(lnY_{it})^2 + \frac{1}{2}\beta_{kk}(lnK_{it})^2 + \frac{1}{2}\beta_{ll}(lnL_{it})^2 +$$

$$\frac{1}{2}\beta_{kl}(lnK_{it} * lnL_{it}) + \beta_{yk}(lnY_{it} * lnK_{it}) + \beta_{yl}(lnY_{it} * lnL_{it}) + \beta_t T + \frac{1}{2}\beta_{tt}T^2 + \beta_{ty}(T * lnY_{it}) + \beta_{tk}(T * lnK_{it}) + \beta_{tl}(T * lnL_{it}) + v_{it} + u_{it} \qquad (2)$$

Where, E_{it} is total energy consumption, Y_{it} is industrial output, K_{it} represents fixed capital, and L_{it} is the number of employees. $D1_i$, $D2_i$, and $D3_i$ are dummy variables for the Eastern China, North Eastern China and Central China regions respectively and i represents 26 Chinese provinces and the period covered, T, is 2006 to 2011. With v_{it} being a white noise error term and u_{it} a non-negative error term capturing the inefficient use of energy, with what their measure of efficiency given by $Eff_{it} = \exp(-\hat{u}_{it})$.[33]

In addition, Lin and Wang (2014) attempt to explain the inefficiency term u_{it} from Equation (2) based on the Battese and Coelli (1995) approach by assuming that it has a systematic component associated with a vector of policy measures (z_{it}) consisting of industry concentration and ownership structure plus a random component (ε_{it}) given by:

$$u_{it} = z_{it}\delta + \varepsilon_{it} \qquad (3)$$

---

[32] Note, Lin and Wang (2014) start by specifying an input distance function, but their estimated equation appears to be consistent with an input requirement specification outlined in the previous section.

[33] Lin and Wang (2014) do not state the software they used for their estimation nor do they appear to have stated the assumption made for the distribution of u_{it},, nor we assume that they have used a truncated normal distribution .

Lin and Wang (2014) conclude that there was an increasing trend in the efficient use of energy in the Chinese iron and steel industry across all regions, although they suggest that there is greater potential for energy savings in the Central and Western regions of China. Moreover, they find that the structural defects in the economic system acts as an impairment to improving the efficient use of energy in the Chinese iron and steel sector. However, Lin and Wang (2014) appear to have ignored the econometric issue related to the presence of unobserved heterogeneity bias that could have been considered given their use of a panel data set. Furthermore, the specification used by Lin and Wang (2014) did not distinction between persistent and transient efficiency.

4.2 Examples of Shephard Energy Distance Function Estimation

An example of research applying the Shephard Energy Distance Function is Zhou et al. (2012) who estimate a SFA PM cross section economy wide model for 21 OECD countries using 2001 data.[34] The log-linear estimated equation is given by:

$$ln(1/E_i) = \beta_0 + \beta_K lnK_i + \beta_L lnL_i + \beta_Y lnY_i + v_i + u_i \qquad (4)$$

Where, E_i is energy consumption, K_i is the capital stock, L_i is the labour force , and Y_i is GDP in country i respectively and v_i is a white noise error term. u_i a non-negative error term capturing Zhou et al.'s measure of energy inefficiency and assumed to be distributed either half-normal or truncated- normal.[35] With what they called the economy-wide energy efficiency index given by $EEI_i = \exp(-\hat{u}_i)$.

The results show a close correlation between those obtained assuming a half-normal distribution and a truncated-normal distribution; the correlation coefficient and the Spearman rank correlation of the EEI estimates both being 0.98.[36] The results for the estimated EEI are illustrated in Figure 4 and it is interesting to compare these with the IEA estimates for energy

---

[34] They actually also estimate a DEA version for comparison.

[35] Zhou at al. (2012) used the FRONTIER 4.1 software for the estimation (Coelli, 1996).

[36] Although the correlation between the EEI estimates from the SFA and DEA were somewhat lower, with the correlation coefficient ranging from 0.65 to 0.73 and the Spearman's rank correlation coefficient ranging from 0.41 to 0.65.

intensity for 2001 illustrated in Figure 5.[37] These show that there appears to be a reasonably close correspondence with the correlation coefficient for energy intensity and the estimated *EEI* being about –0.8 and the Spearman rank correlation coefficient being about –0.9. Perhaps this is not surprising given the limited number of cross section observations for the one year. Moreover, the econometric specification did not make an allowance for taking into account the possible unobserved heterogeneity bias across the countries included in the cross section sample. That said there are still instances where energy intensity appears to overestimate a country's relative *EEI* position. For example, Austria and Finland are ranked 12th and 20th respectively according to energy intensity, whereas the *EEI* measure suggests that they are ranked somewhat higher at 6th/7th and 14th respectively. Whereas, Norway, Spain and Switzerland are ranked 13th, 5th and 1st respectively according to energy intensity, whereas the *EEI* measure suggests that they are ranked somewhat lower at 20th, 9th and 6th/5th respectively.
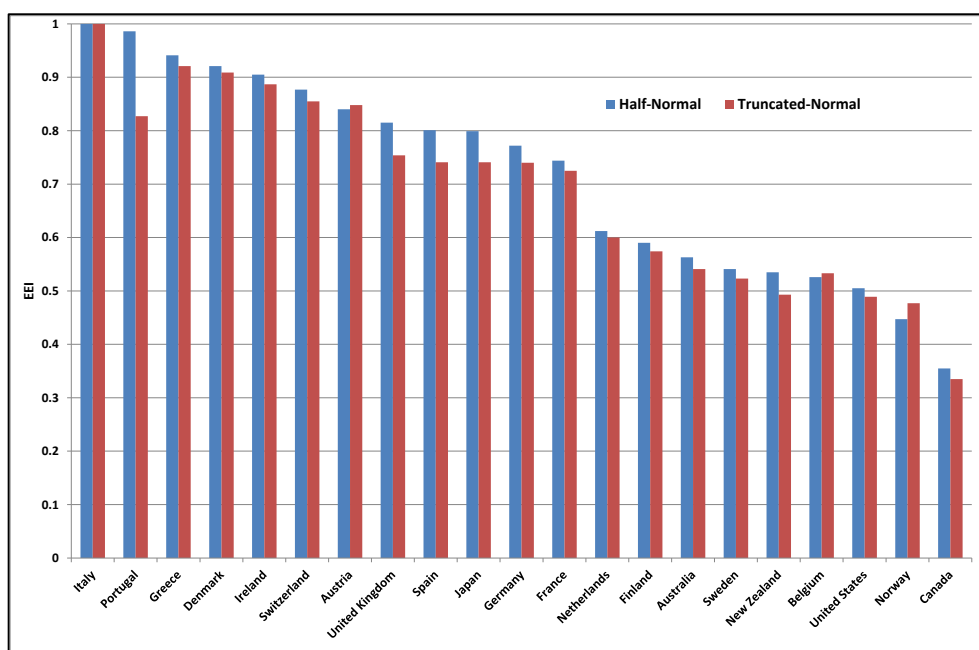
Figure 4: Estimated *EEI* for 2001 from Zhou et al. (2012)
(Ranked according to the Half-Normal estimates)

[37] Energy intensity is measured as Total Final Consumption/GDP, i.e. toe per thousand 2005 USD PPP (IEA, 2014).
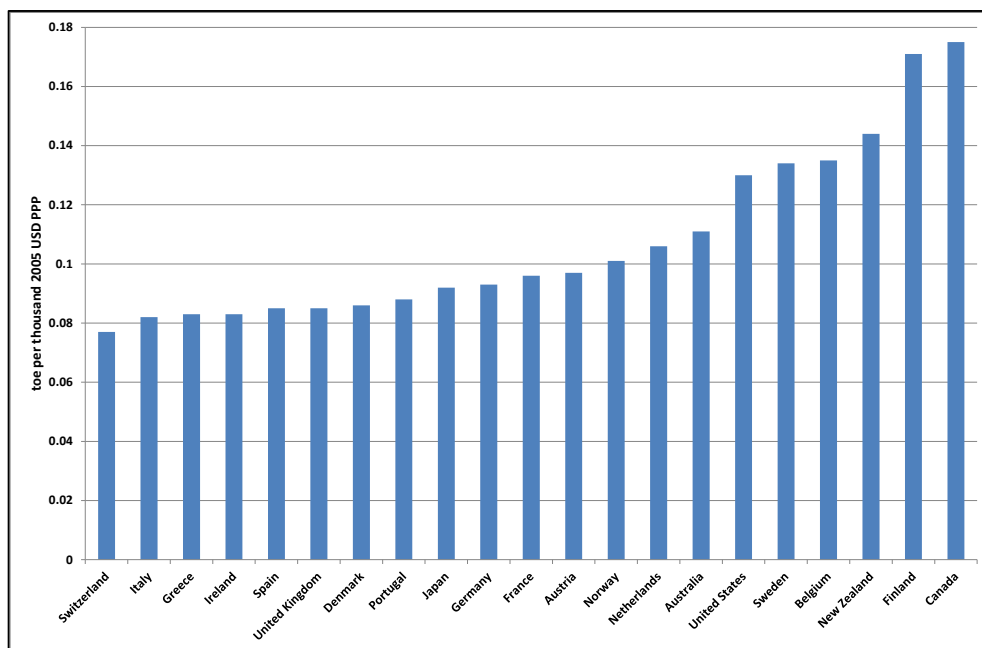
Figure 5: Energy Intensity for 2001 (Source: IEA, 2014)

Lin and Du (2013) also use the Shephard energy distance approach in order to analyse the efficient use of energy across China's 30 administrative regions over the period 1997 to 2010. Lin and Du (2013) also employ a translog specification using energy consumption, GDP, labour and capital, similar to Lin and Wang (2014); however, differ by excluding the technical progress term (T) and utilising $-ln\,E_{it}$ as the dependent variable; effectively equivalent to Zhou et al.'s (2012) dependent variable specification of $ln(1/E_i)$. Although Lin and Du's (2013) data set includes China's 30 administrative regions they actually estimate their model for three groups within the 30 regions based ***not*** on geographical regions (such as the east, central and west provinces) but by using energy intensity[38], stating that the "differences of energy intensity among groups are profound" (p. 532). [39] This however, is arguably a problem since it pre supposes that energy intensity is a good indicator of energy efficiency which is what Lin and Du (2013) are actually attempting to estimate. Nevertheless, they do also estimate what they call a pooled model that includes all the groups but, based on a likelihood ratio test, argue that the frontiers of

---

[38] Their first group included eight regions in the east area (Fujian, Guangdong, Guangxi, Hainan, Shanghai, Zhejiang, Shangdong, Jiangsu) and two regions in the central area (Anhui and Jiangxi). Their second group included three regions in the east area (Beijing, Liaoning, Tianjin), six regions in the central area (Hebei, Henan, Heilongjiang, Hubei, Hunan, and Jilin), and four regions in the west area (Shaanxi, Sichuan, Yunnan, and Chongqing). And their final group included two regions in the central area (Inner Mongolia and Shanxi) and five regions in the west area (Gansu, Guizhou, Ningxia, Qinghai, Xinjiang).

[39] Lin and Du (2013) appear to have used the FRONTIER 4.1 software for the estimation (Coelli, 1996) and used a truncated normal distribution for the inefficiency term u_{it}.

the three groups are not the same. However, Lin and Du (2013) appear to have ignored the issue of potential econometric problem of unobserved heterogeneity bias and perhaps rather than grouping according to energy intensity their pooled model should have included additional variables to address the possible heterogeneity of the groups and/or use appropriate econometric models for panel data as outlined in Section 3. Furthermore, Lin and Du (2013) (and Zhou et al., 2012) did not distinguish between persistent and transient efficiency.[40]

4.3 Examples of Energy Demand Frontier Function Estimation

The final group of examples estimate energy demand frontier functions. Filippini and Hunt (2011) estimated an energy demand frontier function for the whole economy with an unbalanced panel of 29 OECD countries over the period 1978 to 2006, which was extended and updated in Evans et al. (2013) for the period 1978 to 2008. The specification used being:

$$e_{it} = \alpha + \alpha^y y_{it} + \alpha^p p_{it} + \alpha^{pop} pop_{it} + \delta_t D_t + \alpha^C DCOLD_i + \alpha^R DARID_i + \alpha^a a_i +$$
$$\alpha^I ISH_{it} + \alpha^S SSH_{it} + v_{it} + u_{it} \tag{5}$$

where e_{it} is the natural logarithm of aggregate energy consumption (E_{it}), y_{it} is the natural logarithm of GDP (Y_{it}), p_{it} is the natural logarithm of the real price of energy (P_{it}), pop_{it} is the natural logarithm of population (POP_{it}), $DCOLD_i$ is a cold climate dummy variable, $DARID_i$ is a hot climate dummy, a_i is the natural logarithm of the area size of a country (A_i), ISH_{it} is the share of value added of the industrial sector, and SSH_{it} is the share of value added for the service sector. The time variable D_t is a series of time dummy variables.[41] Once again the error term is composed of two independent parts: v_{it} and u_{it}.[42] The level of energy efficiency for each country in each year being given by $EE_{it} = E_{it}^F/E_{it} = exp(-\hat{u}_{it})$.

Given the relatively long time period, Filippini and Hunt's (2011) preferred estimated model and the one presented in Evans et al. (2013) is an energy demand frontier with time dummies to

---

[40] Lin and Du (2013) do however go on to estimate a metafrontier based on linear programming, but this is beyond the scope of this review so is not considered here.

[41] Note, that this model specification does not include the price of capital due to the lack of consistent data.

[42] Filippini and Hunt (2011) and Evans et al. (2013) used the LIMDEP and NLOGIT software for the estimation.

represent the UEDT using a PM.[43] The results show that the estimated *EE* is generally negatively correlated with energy intensity for most countries but with some exceptions. As Filippini and Hunt (2011) and Evans et al. (2013) argue, this is generally to be expected, but if this technique were to be a useful procedure for estimating 'true' energy efficiency then a perfect, or even near perfect, negative correlation would not be expected since all the useful information would be contained in the standard energy intensity measure. This is confirmed, given that in Evans (2013) the average correlation coefficient between the estimated *EE* and energy intensity across all countries over the whole estimation period is −0.48. Moreover, within this, there is a relatively high negative correlation for some countries, such as the Czech Republic, Denmark, Germany, Greece, Hungary, Ireland, Luxembourg, the Netherlands, New Zealand, Poland, Portugal, Slovak Republic, the UK and the USA, whereas for some countries the (negative) correlation is somewhat less, such as Canada, Finland, France, Korea, Norway, Spain and Sweden. Furthermore, for Australia, Austria, Belgium, Italy, Japan, Mexico, Switzerland and Turkey, there appears to be a positive relationship between energy intensity and estimated energy efficiency. This suggests that for some countries energy intensity is a reasonable proxy for 'true' energy efficiency, whereas for others it is a very poor proxy.

---
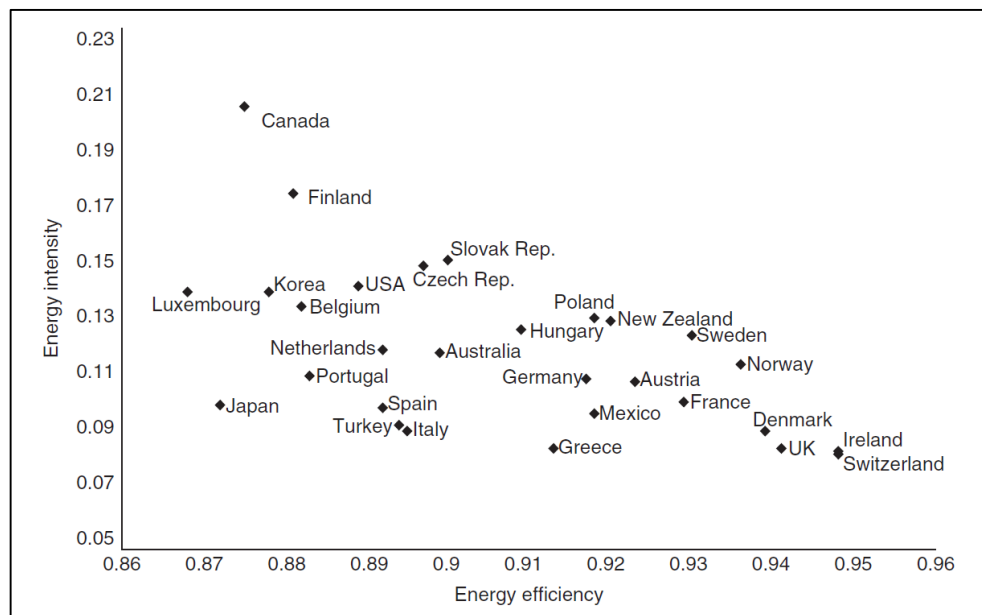
[43] Filippini and Hunt (2011) also estimated a PM with a time trend and a TREM with a time trend for comparison. The PM with the time dummies was preferred to the PM with the time trend given its non-linear representation of the UEDT. For the TREM, they argue that given certain sources of energy inefficiency that result in time-invariant excess energy consumption, the estimates from the model provide imprecise estimates resulting in overestimated levels of energy efficiency. However, recent developments as discussed in Section 3 suggest that estimates from the TREM might in fact represent transient energy efficiency.

**Figure 6: Estimated EE compared to Energy Intensity 2003 to 2007
        from Evans et al. (2013)**

Figure 6 focuses on the Evans et al. (2013) results for the period 2003 to 2007 and shows that although there would appear to be a generally negative relationship between the rankings of estimated EE and energy intensity, there is not a one- to- one correspondence. For example, according to the energy intensity measure, Italy, Turkey and Japan are ranked 6th, 7th and 10th respectively whereas they are estimated to be 18th, 19th and 28th respectively according to the estimated EE measure; thus for these countries the simple energy intensity ratio would appear to overestimate their 'true' relative energy efficiency position. On the other hand, according the energy intensity ratio Sweden, New Zealand, Poland and the Slovak Republic are ranked 18th, 20th, 21st and 27th respectively, whereas they are estimated to be 6th, 9th, 11th and 15th respectively according to the estimated EE measure; thus for these countries energy intensity would appear to underestimate their 'true' relative energy efficiency position.

Filippini and Hunt (2012) and Filippini et al. (2014) build on the approach suggested by Filippini and Hunt (2011)[44] but use a shorter time period and attempt to further control for unobserved heterogeneity by estimating models with a Mundlak adjustment (Mundlak, 1978). Filippini and Hunt (2012) estimate a US frontier residential aggregate energy demand function using panel data for 48 'states' over the period 1995 to 2007 using the REM with the Mundlak adjustment. [45] They find that energy intensity is not necessarily a good indicator of their estimated levels of energy efficiency.

Filippini et al. (2014) also prefer a Mundlak corrected REM, but augment the energy demand frontier with an equation to assess the impact of various energy efficiency policies on the estimated EE. Filippini et al. (2014) use an unbalanced panel data set for a sample of 27 EU member states over the period 1996 to 2009 and estimate a log-linear residential equation given by:

$$ln\,ED_{it} = a + b_{PE}lnPE_{it} + b_Y lnY_{it} + b_{POP}lnPOP_{it} + b_{DSIZE}lnDSIZE_{it} + b_{HDD}lnHDD_{it} +$$
$$b_{HOT}lnHOT_i + b_t t + v_{it} + u_{it} \tag{6}$$

---

[44] Weyman-Jones et al. (2005) is another example of research attempting to estimate the efficient use of energy based on the Energy Demand Frontier approach, but for Portuguese electricity consumption.

[45] Filippini and Hunt (2012) used the LIMDEP and NLOGIT software for the estimation.

where ED_{it} represents final residential energy consumption, PE_{it} the real energy price, Y_{it} real income, POP_{it} population, $DSIZE_{it}$ the average size of a dwelling, HDD_{it} the number of heating degree days, HOT_i a hot climate dummy variable, and t a time trend to proxy a UEDT. With v_{it}, u_{it} and UEE_{it} defined as in Filippini and Hunt (2011) and Evans et al. (2013). Filippini et al. (2014) estimate their preferred model using the Battese and Coelli (1995) approach but with a Mundlak adjustment to their model, like Filippini and Hunt (2012). Furthermore, in a similar way to Lin and Wang (2014), Filippini et al. (2014) assume that their inefficiency term u_{it} has a systematic component associated with a vector of policy measures ($\mathbf{Z}_{it}$) and a random component (ε_{it}) given by:

$$u_{it} = \eta' Z_{it} + \varepsilon_{it} \tag{7}$$

Several energy-efficiency policy measures are considered within $\mathbf{Z}_{it}$, such as performance standards of buildings and heating systems, performance standards of electrical appliances, financial incentives and informative measures. The model was therefore estimated in a single stage by the maximum likelihood procedure. [46] Similar to previous results discussed above, the estimated EE scores from the preferred model were found to be poorly correlated with energy intensity, measured by the average residential energy demand and with the average consumption per dwelling; again suggesting that energy intensity is a poor proxy or the 'true' level of energy efficiency.

The results from Filippini et al.'s (2014) augmented equation show that a number of so-called energy efficiency instruments influence the level of u_{it} of the residential sector. In particular, financial incentives have an important influence on reducing EU states' inefficient use of energy. In addition, the results suggest that performance standards of buildings, heating systems and appliances also contribute to reducing the inefficient use of energy. However, the results suggest that informative measures do not reduce energy inefficiency (and in fact, counter intuitively, suggest that they increase the level of energy inefficiency).

It is worth highlighting that the energy demand frontier models used by Filippini and Hunt (2011), Evans et al. (2013), Filippini and Hunt (2012), and Filippini et al. (2014), given the

---

[46] Filippini et al. (2014) used the LIMDEP and NLOGIT software for the estimation.

upstanding at the time, did not make a distinction between permanent and transient energy efficiency, as discussed in Section 3 above However, more recently Filippini and Hunt (2015) attempted to measure persistent and transient EE for the whole economy of 49 states in the US using a stochastic frontier energy demand approach using two separate econometric models (as discussed in Section 3).[47] They use data for 49 states over the period 1995 to 2009 using two panel data models: the Mundlak version of the REM (which estimates the persistent part of energy efficiency) and the Mundlak version of the TREM (which estimates the transient part of energy efficiency). Their basic log-linear equation is given by:

$$e_{it} = \alpha + \alpha^p p_{it} + \alpha^y y_{it} + \alpha^{pop} pop_{it} + \alpha^{hs} hs_{it} + \alpha^{hdd} hdd_{it} + \alpha^{cdd} cdd_{it} + \alpha^{SHI} SHI_{it} +$$
$$\alpha^{SHS} SHS_{it} + \alpha^a a_i + \alpha^t t + v_{it} + u_{it} \qquad\qquad (8)$$

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 GDP (Y_{it}), pop_{it} is the natural logarithm of population (POP_{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}), hs_{it} is the natural logarithm of the household size (HS_{it}), a_i is the natural logarithm of the area size (A_i), SHI_{it} is the share of value added of the industrial sector, and SHS_{it} is the share of value added for the service sector, and t is a time trend that proxies the UEDT. Again, v_{it}, u_{it}, and estimated EE_{it} are defined as above.

The results obtained by Filippini and Hunt (2015) suggest that the estimated average persistent EE is greater than the transient part, with the variation in the estimated transient EE being somewhat lower than the variation in the estimated persistent EE. Furthermore, consistent with above, their results suggest that energy intensity is not a good indicator of the 'true' level of energy efficiency, whereas, by controlling for a range of economic and other factors, their measure of persistent EE provides better a guide to the 'true' efficiency; illustrated in Figure 8. Moreover, the correlation coefficients between energy intensity and the estimated average EE measure from the Mundlak adjusted REM and the Mundlak adjusted TREM are -0.46 and -0.21 respectively. In addition, there is only a weak correlation between the rankings, with the

---

[47] But not the combined approach recently developed by Filippini and Greene (2015).

Spearman rank correlation coefficients between energy intensity and the average *EE* measure from the two models being 0.18 and 0.21 respectively.

Therefore, again consistent with the studies reviewed above, the Filippini and Hunt (2015) estimates suggest that although for some states energy intensity might give a reasonable indication of a state's relative estimated *EE* this is not the case for all states, California being a prime example. As Filippini and Hunt (2015) point out, this result is not consistent with the conventional wisdom of energy efficiency policymakers and professionals who generally regard California as being a highly energy efficient state as well as a number of research papers such as Howrowitz (2007) and Sudarsham (2013). However, the view is normally based on energy intensity (or just electricity intensity) so a direct comparison with the analysis here is difficult if not impossible given the main idea of the estimated *EE* measure is that analysis based on energy intensity is potentially biased and misleading for policymakers. Thus, Filippini and Hunt (2015) argue that their results do not implicitly disagree with some of the previous research (such as Howrowitz, 2007) who suggest that California's energy efficiency programs have contributed to reducing the state's electricity intensity, just that there is still more to be done in order for California to increase its energy efficiency and move closer to the energy demand efficient frontier.
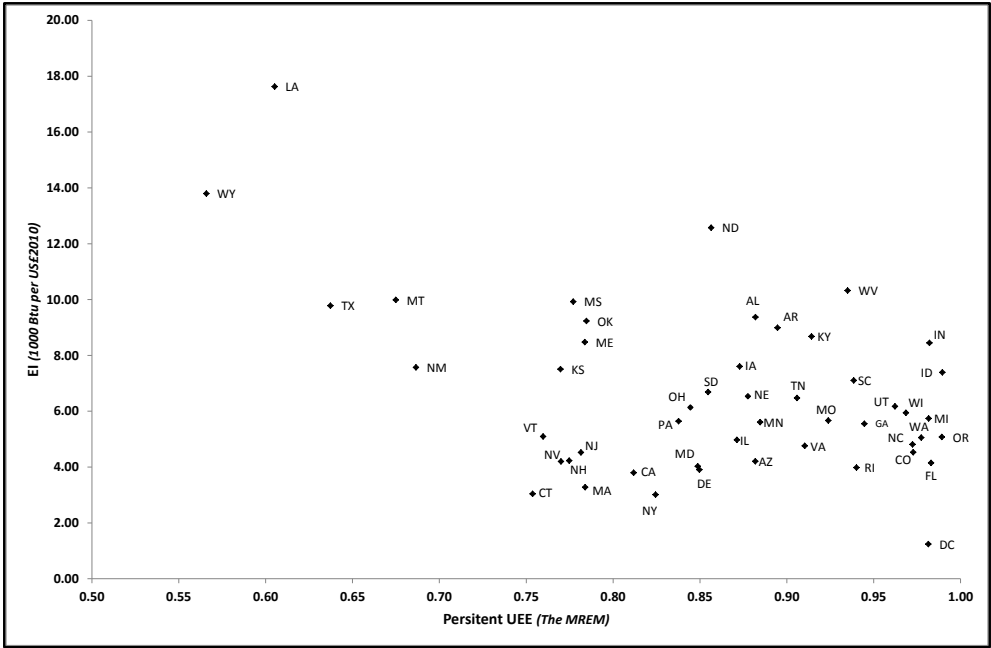
Figure 7: Scatter diagram of average energy intensity and estimated persistent energy efficiency (1995-2009) from Filippini and Hunt (2015)

5. Summary and Conclusion

This paper attempts to present a clear theoretical background for measuring the level of energy efficiency based on economic foundations. It is, as far as is known, one of the first attempts to provide a discussion on the relation between the concepts of productive efficiency and energy efficiency, to discuss the theoretical approaches that can be used and to present parametric econometric methods that can be applied in measuring the level of energy efficiency. Furthermore, we are of the view that definition of energy efficiency based on an economics perspective outlined in this paper and resultant estimated economic based energy efficiency indicators are better than relying on simple energy intensity indicators.

The theoretical section argues that the definition of energy efficiency should be based on the non-radial notion of input specific technical efficiency measure introduced by Kopp (1981). This measure, in terms of the input energy, is defined as the ratio of minimum feasible amount of energy to the observed use of energy – conditional on the production technology and the observed levels of outputs and other inputs. Moreover, a variant of this measure was proposed by Filippini and Hunt (2011) whereby the function indicates the minimum amount of energy that is necessary to produce a given level of outputs given technology and input prices. In this context, actual energy demand differs from the estimated energy frontier demand due to the presence of both allocative and technical inefficiency.

Furthermore, it is shown that from an empirical perspective, there are three approaches that would appear to be more appropriate for estimating parametric methods of the non-radial notion of energy efficiency: the energy requirement function; the energy demand function; and a shepherd energy distance function. From these approaches, we believe that the estimation of an energy demand function is an interesting approach and perhaps more appealing than the other two options because it considers both allocative and technical efficiency. However, it is recognised that this approach has some limitations that should be considered in any empirical application and when interpreting results.

Moreover, the paper illustrates the recent developments in the econometric empirical techniques used to estimate such models from the PM and the TREM to the GTREM. Some of these produce time invariant values of the level of efficiency (persistent efficiency), others models estimate time variant values (transient efficiency), whereas the more recent GTREM is a technique that

allows for both components of efficiency concurrently. Moreover, we believe that this distinction in the type of efficiency is very relevant and will, in all probability, gain importance in future research on the estimation of the efficient use of energy.

Although this does not pretend to be a definite view on the matter, it is hoped that this review will help stimulate further debate and the development of such models, such as the combined estimation of both transitory and persistent energy efficiency. It is therefore hoped that it will be seen as a step in the right direction rather than just relying on the poor proxy of energy intensity.

References

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.

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

Bhattacharyya S.C. (2011), *Energy Economics Concepts, Issues, Markets and Governance*, Springer-Verlag.,

Battese, G.E. and T. J. Coelli (1995) 'A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data', *Empirical Economics,* **20**, 325–332.

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

Chambers, R.G. (1988). *Applied Production Analysis*. Cambridge University Press.

Coelli, T. J. (1996) 'A guide to FRONTIER 4.1: a computer program for stochastic frontier production and cost function estimation', *CPEA working papers (No. 7/96). Department of Econometrics, University of New England*.

Colombi, R., Kumbhakar, S., Martini, G., & Vittadini, G. (2014) 'Closed-skew normality in stocastic frontiers with individual effects and long/short-run efficiency', *Journal of Productivity Analysis*. Published online

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

Evans, J., Filippini, M. & Hunt, L. C. (2013) 'The contribution of energy efficiency towards meeting CO2 targets', Chapter 8 in Fouquet, R. (ed) *Handbook of Energy and Climate Change,* Cheltenham, UK and Northampton, MA, USA: Edward Elgar, 175-223.

Farrell, M. (1957) 'The measurement of productive efficiency', *Journal of the Royal Statistical Society, Series A, General*, **120**, 253–81.

Farsi, M. and M. Filippini (2009) 'Efficiency Measurement in the Electricity and Gas Distribution Sectors', Chapter 25 in J. Evans and L. C. Hunt (eds) *International Handbook on the Economics of Energy*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar, 598-623.

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

Filippini, M. and W. Greene, W. (2015) 'Persistent and Transient Productive Inefficiency: A Maximum Simulated Likelihood Approach', *Journal of Productivity Analysis, forthcoming*.

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**, 59–80.

Filippini M. and L. C. Hunt (2012) 'US Residential Energy Demand and Energy Efficiency: A Stochastic Demand Frontier Approach', *Energy Economics,* **34**, 1484–1491.

Filippini M. and L. C. Hunt (2015) 'Underlying Energy Efficiency in the US', *Surrey Energy Economics Discussion paper Series*, SEEDS150. Surrey Energy Economics Centre (SEEC), University of Surrey, Guildford, UK.

Filippini M., L. C. Hunt and J. Zorić (2014) 'Impact of energy policy instruments on the estimated level of underlying energy efficiency in the EU residential sector', *Energy Policy,* **69**, 73–81.

Gillingham, K., R.G. Newell and K. Palmer (2009), 'Energy efficiency economics and policy', *Annual Review of Resource Economics*, **1**, 597–619.

Greene, W. H. (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.

Greene, W. H. (2008) The econometric approach to efficiency analysis, Chapter 2 in *The Measurement of Productive Efficiency and Productivity Growth*, H. O. Fried, C. A. K. Lovell, and S. S. Schmidt, (eds.). Oxford University Press, 92–250.

Guan, Z., S. C. Kumbhakar, R. J. Myers, and A. O. Lansink,(2009) 'Measuring excess capital capacity in agricultural production', *The American Journal of Agricultural Economics*, 91, 765–776,

Hu, J. and S. Wang (2006): Total-factor energy efficiency of regions in China. *Energy Policy* 34, 3206-3217.

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: L. C. Hunt, (ed.), *Energy in a Competitive Market: Essays in Honour of Colin Robinson*, Edward Elgar, Cheltenham, UK and Northampton, MA, USA, 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/Textbase/publications/free_new_Desc.asp?PUBS_ID=2127

IEA (2014): Energy Balances of OECD Countries (Edition: 2014), International Energy Agency, Mimas, University of Manchester. DOI: http://dx.doi.org/10.5257/iea/ebo/2014

Jaffe, A. and R. Stavins (1994) 'The energy efficiency gap: What does it mean?', *Energy Policy,* **22**, 804–810.

Karagiannis, G, V. Tzouvelekas and A. Xepapadeas (2003) 'Measuring irrigation water efficiency with a stochastic production frontier: An application to Greek out-of-season vegetable cultivation', *Environmental and Resource Economics*, 26, 57–72.

Kopp, R.J. (1981) The Measurement of Productive Efficiency: A Reconsideration, *The Quarterly Journal of Economics*, 96, pp. 477-503.

Kumbhakar, S.C., and L. Hjalmarsson (1995). 'Labor-Use Efficiency in Swedish Social Insurance Offices' *Journal of Applied Econometrics* **10**, 33-47.

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

Kumbhakar, S. C. and Tsionas, E. G. (2012). 'Firm Heterogeneity, Persistent and transient Technical Inefficiency: A generalized True Random Effects model' *Journal of Applied Econometrics*, *29*, 110–132.

Lin, B. and X. Wang (2014) 'Technology gap and China's regional energy efficiency: A parametric metafrontier approach', *Energy Policy,* **72**, 87–96.

Mundlak, Y. (1978) 'On the Pooling of Time Series and Cross-Section Data', *Econometrica*, **46**, 69–85.

Murillo-Zamorano, L. R. (2004). Economic Efficiency and Frontier Techniques, *Journal of Economic Surveys*, vol. 18, no. 1, pp. 33-77.

Patterson, M. G. (1996) 'What is energy efficiency? Concepts, indicators and methodological issues,' *Energy Policy*, 24 : 377–390.

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.

Reinhard S, C. A. K. Lovell and G. J. Thijssen (1999). 'Econometric estimation of technical and environmental efficiency: An application to Dutch dairy farms', *American Journal of Agricultural Economics,* **81**, 44–60.

Schmidt, P. and C. A. K. Lovell (1979) 'Estimating technical and allocative inefficiency relative to stochastic production and cost frontiers', *Journal of Econometrics*, 9, 343–66.

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

Sorrell, S. (2009), 'The rebound effect: definition and estimation', in J. Evans and L.C. Hunt (eds), *International Handbook on the Economics of Energy*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar, 199–233.

Wei, C., J. Ni, and M. Shen (2009): Empirical analysis of provincial energy efficiency in China. China & World Economy, 88-103, **17**(5), 88-103.

Weyman-Jones, T., J. M. Boucinha and C. F Inácio (2015) "Measuring electric energy efficiency in Portuguese households", *Management of Environmental Quality: An International Journal*, **26**(3), 407 – 422.

Zhou, P. and 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 D. Q. Zhou (2012) 'Measuring economy-wide energy efficiency performance: A parametric frontier approach', **Applied Energy, 90**, 196– 200.

Working Papers of the Center of Economic Research at ETH Zurich

(PDF-files of the Working Papers can be downloaded at www.cer.ethz.ch/research).