Università
della
Svizzera
italiana

Faculty
of Economics

IdEP Economic Papers

2016 / 03

M. Filippini, W. Greene, G. Masiero

Persistent and transient productive inefficiency in a regulated industry: electricity distribution in New Zealand

Persistent and transient productive inefficiency in a regulated industry: electricity distribution in New Zealand

M. Filippini * W. Greene † G. Masiero ‡

February 2016

Abstract

The productive efficiency of a firm can be decomposed into two parts, one persistent and one transient. So far, most of the cost efficiency studies estimated frontier models that provide either the transient or the persistent part of productive efficiency. This distinction seems to be appealing also for regulators. During the last decades, public utilities such as water and electricity have witnessed a wave of regulatory reforms aimed at improving efficiency through incentive regulation. Most of these regulation schemes use *benchmarking*, namely measuring companies' efficiency and rewarding them accordingly. The purpose of this study is to assess the level of persistent and transient efficiency in an electricity sector and to investigate their implications under price cap regulation. Using a theoretical model, we show that an imperfectly informed regulator may not disentangle the two parts of the cost efficiency; therefore, they may fail in setting optimal efficiency targets. The introduction of minimum quality standards may not offer a valid solution. To provide evidence we use data on 28 New Zealand electricity distribution companies between 1996 and 2011. We estimate a total cost function using three stochastic frontier models for panel data. We start with the random effects model (RE) proposed by Pitt and Lee (1981) that provides information on the persistent part of the cost efficiency. Then, we apply the true random effects model (TRE) proposed by Greene (2005a, 2005b) that provides information on the transient part. Finally, we use the generalized true random effects model (GTRE) that allows for the simultaneous estimation of both transient and persistent efficiency. We find weak evidence that persistent efficiency is associated to higher quality, and wrong efficiency targets are associated to lower quality compliance.

Keywords: cost efficiency, regulation, persistent and transient productive efficiency, electricity distribution.

JEL classification: C1, C23, D24.

*Institute of Economics (IdEP), Università della Svizzera italiana (USI); Department of Management, Technology and Economics, ETH Zurich, Switzerland. Corresponding author. E-mail address: massimo.filippini@usi.ch. We are indebted to Heinke Wetzel for providing the dataset used in the empirical part of the paper.

†Department of Economics, Stern School of Business, New York University, USA.

‡Department of Management, Information and Production Engineering, University of Bergamo, Italy; Institute of Economics (IdEP), Università della Svizzera italiana (USI), Switzerland.

1 Introduction

During the last twenty years, several countries around the world have introduced reforms in public utility sectors, such as water, electricity and telecommunications. Regarding electricity, two key elements of these reforms are the introduction of competition in the supply and generation of electricity, and the introduction of new regulation methods in the transmission and distribution of electricity considered as natural monopolies. The new methods apply the incentive regulation theory (Laffont and Tirole, 1993). They provide incentives for productive efficiency by compensating the company with its savings. The main categories of incentive-based schemes used for electricity utilities are price or revenue caps, sliding-scale rate of return, partial cost adjustment, menu of contracts, and yardstick regulation.[1] Several of these methods used by regulation authorities make use of information on the level of productive efficiency of an electricity distribution company, i.e. technical and cost efficiency. For instance, the New Zealand regulation authority for the electricity distribution sector has adopted a price-cap regulation scheme.

As discussed in Filippini and Greene (2015), the level of productive efficiency of a firm can be decomposed into two parts, one persistent and one transient. The presence of structural problems in the organization of the production process or systematic shortfalls in managerial capabilities can generate the persistent part. Conversely, the presence of non-systematic management problems in the short term determines the transient part. For the regulator it is crucial to distinguish between these two types of inefficiency, and to choose measurement methods of productive efficiency that provide information on both persistent and transient components.

So far none of the regulation authorities around the world that are using information on the level of productive efficiency of electricity distribution companies in the regulation process make a distinction between persistent and transient inefficiency. The reasons are twofold. First, only recently have some scholars introduced this distinction. Secondly, the empirical measurement of these two components of the level of productive efficiency is still in a development phase. However, the literature on the estimation of the productive efficiency that does

[1] See Joskow and Schmalensee (1986) for a review of regulation models.

not distinguish the two components is abundant.[2]

A few studies analyze the level of cost efficiency of the New Zealand electricity distribution companies. None of these studies makes any distinction of the two components of productive efficiency, i.e. transient and persistent. Scully (1999) analyzes the impact of privatization on the level of cost efficiency using a sample of electricity distribution companies operating in New Zealand. The author proposes the estimation of a translog total cost function using data from 1982 to 1994. The results show that privatization produces substantial cost reductions. Filippini and Wetzel (2014) assess the cost efficiency of 28 New Zealand electricity distribution companies for the period between 1996 and 2011. The authors estimate a total cost function and a variable cost function using a stochastic frontier model with panel data. The goal of this study is to evaluate the impact of ownership unbundling on the level of cost efficiency. The results suggest that ownership separation of electricity generation and retail operations from the distribution network have positive effects on the level of cost efficiency.[3]

The goals of this paper are twofold. First, it sketches a theoretical model that shows the importance of the distinction between persistent and transient inefficiency in the application of a price-cap regulation method. Second, the paper estimates the level of cost efficiency for a sample of 28 New Zealand electricity distribution companies by making the distinction between persistent and transient levels of efficiency. The paper intends to contribute to the literature in two ways. First, the theoretical model illustrates how to use information on persistent and transient inefficiency effectively in a price-cap regulation setting. Second, the paper provides one of the first empirical analyses to show the presence of persistent and transient inefficiency using a novel econometric approach introduced by Filippini and Greene (2015). The results support the importance of the distinction of productive efficiency into a persistent and a transient part as depicted in the theoretical model.

We organize the paper as follows. Section 2 sketches a theoretical model that investigates how regulation may fail when persistent and transient inefficiency are ignored. Section 3 introduces the cost model specification and the estimation

[2]See Ramos-Real (2005) for a review of part of these studies.

[3]Nillesen and Pollitt (2011) also investigate the effect of ownership unbundling on electricity distribution companies but their analysis focuses on the level of cost rather than the level cost efficiency, and only considers variable costs.

approaches, while Section 4 describes the data. In Section 5, we present and discuss the estimation results. Section 6 summarizes and concludes.

2 Persistent and transient inefficiency in a regulated industry: a theoretical approach

To understand the implications of transient and persistent inefficiency of electricity distribution companies in New Zealand, we sketch a model where firms maximize the current value of future profits under price cap regulation. In the market there are N identical firms acting as local monopolies.[4] Each firm chooses price (p_t) and service quality (q_t) in each period t as well as the level of managerial effort (e_t). Managerial effort allows obtaining efficiency gains only partially exploitable in terms of cost reductions because of persistent inefficiency. We show that the regulator cannot achieve optimal efficiency targets if imperfectly informed on persistent inefficiency. Regulation failure may lead to postponed expenditures, which worsens service quality, or increases in monopoly rents. In addition, higher pressure to meet current minimum quality standards may undermine a firm's compliance and also result in delayed expenditures and poorer quality in the future.[5]

Consider first the following demand for electricity distribution services faced by each firm:

$$s(p_t, q_t) = q_t\left(\theta - p_t\right), \tag{1}$$

where $q_t \in (0, +\infty)$ is service quality, p_t is unit price, and θ is a parameter indicating the reservation price for a unit of quality.

The following equation describes total costs at t:

$$c(e_t, e_{t-j}, p_t, q_t) = \beta + \gamma s(p_t, q_t) - E_T\left(e_t\right) - E_P\left(e_{t-j}\right) + f(q_t) + g(e_t), \tag{2}$$

where β is a cost that depends on the size of the distribution network and γ is the unit cost of electricity services ($\gamma < \theta$). Both quality and managerial effort are

[4] In practice, firms are single-product monopolists. For a model of price cap regulation with multi-market monopolists when the costs of serving different markets vary, see for instance Cowan (1997a).

[5] Recently, Di Giorgio et al. (2015) proposed a theoretical approach to separate structural (or institutional) inefficiency from managerial inefficiency in public and private nursing homes. The model applies to a different regulatory setting - global budget instead of price cap regulation - and does not elaborate on the implications for the regulatory mechanism.

costly for the firm. Therefore, $f(q_t) = q_t^2/2$ is the cost of quality and $g(e_t) = e_t^2/2$ is the cost of managerial effort.[6] We assume increasing marginal costs of quality and effort. Managerial effort allows to reach efficiency gains that reduce costs in the current period by $E_T(e_t) = (1-\alpha)e_t$ and $E_P(e_{t-j}) = \alpha e_{t-j}$. Since part of the inefficiency persists over time, only a fraction $(1-\alpha)$ of current managerial effort translates immediately into efficiency gains. The remaining fraction (α) generates efficiency gains only in the future, at time $t+j$. Therefore, the lagged efficiency effort $E_P(e_{t-j})$ also reduces total costs in period t. Cost persistency may arise for several reasons. For instance, management habits may lead to inertia, which prevents improving tasks or solving problems immediately. In addition, environmental and social constraints related to shareholders' preferences, access to inputs, or the fulfillment of legal rules may affect the timing of efficiency gains. Finally, unions' bargaining power may succeed in postponing the achievement of efficiency targets. Therefore, in all these cases the effects of managerial effort on cost reduction may be delayed.

The firm maximizes the current value of future economic profits in each period. Therefore, the firm's intertemporal profits at time t can be written as

$$V_t = \sum_{t=k}^{\infty} \delta^{t-k} \pi_t(e_t, e_{t-j}, p_t, q_t), \tag{3}$$

where

$$\pi_t(e_t, e_{t-j}, p_t, q_t) = p_t s(p_t, q_t) - c(e_t, e_{t-j}, p_t, q_t), \tag{4}$$

and $\delta \leq 1$ is the discount factor for future profits.

For comparison purposes, we first analyze the case where the firm is not regulated. When choosing price, quality and efficiency effort each firm takes into account the effects not only on its current period profits but also on its demand and costs in the following periods. This dependence needs to be taken into account when solving the model for the equilibrium levels of price, quality and effort. Profits in period t depend upon efficiency effort in period $t-j$. In addition, the value function represented by the flow of all future profits depends on all future levels of price, quality and efficiency effort. In equilibrium the firm selects price, quality and efficiency effort that maximize its intertemporal

[6] For instance, one can think to the cost of remunerating the performance of the manager through an increase in the wage.

profit given its subsequent choices of price, quality and efficiency effort. Because efficiency effort affects profits in the subsequent period and expected profits are the sum of concave functions in price, quality and efficiency effort, we can write the following first-order conditions for the firm using Eqs. (1)-(2) and (3)-(4) as:

$$\begin{aligned} &\frac{\partial V_t}{\partial p_t} = \frac{\partial \pi_t\left(e_t, e^*_{t-j}, \boldsymbol{p_t}, q_t\right)}{\partial p_t} = q_t(\theta+\gamma) - 2q_t p_t = 0 \\ &\frac{\partial V_t}{\partial e_t} = \frac{\partial \pi_t\left(\boldsymbol{e_t}, e^*_{t-j}, p_t, q_t\right)}{\partial e_t} + \delta^j \frac{\partial \pi_{t+j}\left(e^*_{t+j}, \boldsymbol{e_t}, p^*_{t+j}, q^*_{t+j}\right)}{\partial e_t} = \\ &\quad = -e_t + (1-\alpha) + \delta^j \alpha = 0 \\ &\frac{\partial V_t}{\partial q_t} = \frac{\partial \pi_t\left(e_t, e^*_{t-j}, p_t, \boldsymbol{q_t}\right)}{\partial q_t} = (p_t - \gamma)(\theta - p_t) - q_t = 0. \end{aligned} \tag{5}$$

Therefore, an equilibrium is defined by:[7]

$$\begin{aligned} &p^*_t = \frac{\theta+\gamma}{2} \\ &e^*_t = 1 - \alpha\left(1-\delta^j\right) \\ &q^*_t = \frac{(\theta-\gamma)^2}{4}. \end{aligned} \tag{6}$$

Note that the equilibrium level of efficiency effort is less than 1 since the expected benefits of effort in terms of future cost reduction are affected by cost persistency (α). However, in the extreme case of $\delta = 1$, i.e. when future costs are not discounted, persistency would not influence the choice of efficiency effort. This is because the impact of efficiency effort on costs in different time periods is the same.

Proposition 1 *Firm's efficiency effort decreases with persistent efficiency (α), provided that the discount factor (δ) on future earnings is less than 1.*

2.1 Price cap regulation

We now assume that the energy authority decides to introduce a price cap regulation, starting at $t+1$, and summarized by the following rule:

$$\frac{p_{t+1} s(p^*_t, q^*_t)}{p^*_t s(p^*_t, q^*_t)} \leq \frac{CPI_{t+1}}{CPI_t} - X, \tag{7}$$

where CPI_{t+1} and CPI_t are Consumer Price Indexes in period $t+1$ and t, respectively.[8] X is the expected efficiency gain based on past performance or

[7] The second-order sufficient conditions are: $\frac{\partial^2 V^*_t}{\partial p^2_t} = -2q^*_t < 0$, $\frac{\partial^2 V^*_t}{\partial e^2_t} = \frac{\partial^2 V^*_t}{\partial q^2_t} = -1 < 0$, $\frac{\partial^2 V^*_t}{\partial p^2_t}\frac{\partial^2 V^*_t}{\partial e^2_t} - \frac{\partial^2 V^*_t}{\partial p_t \partial e_t} = 2q^*_t > 0$, $\frac{\partial^2 V^*_t}{\partial p^2_t}\frac{\partial^2 V^*_t}{\partial q^2_t} - \frac{\partial^2 V^*_t}{\partial p_t \partial q_t} = 2\left(q^*_t + p^*_t\right) - (\theta+\gamma) > 0$, $\frac{\partial^2 V^*_t}{\partial e^2_t}\frac{\partial^2 V^*_t}{\partial q^2_t} - \frac{\partial^2 V^*_t}{\partial e_t \partial q_t} = 1 > 0$. These are satisfied for any $\gamma < \theta$ provided that $\theta \geq 2$.

[8] We consider a standard rule, though alternative price-cap schemes are possible. See Cowan (1997b) for a comparison of different price-cap schemes in terms of allocative efficiency.

average performance of other firms in the market.[9] Let us assume that prices do not inflate, i.e. $CPI_{t+1} = CPI_t$. Therefore, the price cap rule in Eq. (7) can be simplified as:

$$p_{t+1} \leq p_t^* (1 - X) . \tag{8}$$

2.1.1 Full information on efficiency structure

The least realistic scenario hypothesizes that the energy authority has full information about the efficiency structure.[10] In other words, the regulator is fully aware that part of the efficiency effort cannot translate into efficiency gains immediately because of persistency. We start our analysis from the full information case since this is the ideal scenario and can be used as a benchmark. To simplify the analysis we now set unit costs at zero, therefore $\gamma = 0$.[11]

The firm's maximization problem set by Eq. (3) and subject to the price cap constraint defined by Eq. (8) leads to the following first-order conditions:[12]

$$\begin{array}{l} \frac{\partial V_{t+1}}{\partial p_{t+1}} = q_{t+1}\theta - 2q_{t+1}p_{t+1} - \lambda + \delta\lambda (1 - X) = 0 \\ \frac{\partial V_{t+1}}{\partial e_{t+1}} = -e_{t+1} + (1 - \alpha) + \delta^j \alpha = 0 \\ \frac{\partial V_{t+1}}{\partial q_{t+1}} = p_{t+1} (\theta - p_{t+1}) - q_{t+1} = 0 \\ \frac{\partial V_{t+1}}{\partial \lambda} = p_{t+1} - p_t^* (1 - X) = 0, \end{array} \tag{9}$$

where λ is the slack variable. The solution to the constrained maximization problem is:

$$\begin{array}{l} p_{t+1}^* = p_t^* (1 - X) = \frac{\theta}{2} (1 - X) \\ e_{t+1}^* = 1 - \alpha \left(1 - \delta^j\right) \\ q_{t+1}^* = p_t^* (1 - X) [\theta - p_t^* (1 - X)] = \frac{\theta^2}{4} (1 - X)^2 \\ \lambda^* = \frac{q_{t+1}^* \theta - 2q_{t+1}^* p_{t+1}^*}{[1 - \delta(1 - X)]} = \frac{\theta^3}{4} \frac{(1 - X)^2 X}{[1 - \delta(1 - X)]} . \end{array} \tag{10}$$

[9]See Bernstein and Sappington (1999) for a review of the relevant basic principles to determine the X factor.

[10]The problem of information acquisition under price cap regulation has been widely investigated in the literature. See Iossa and Stroffolini (2002) among others.

[11]From Eq. (6) we see that the marginal cost of electricity distribution affects price and quality in equilibrium. However, the equilibrium level of efficiency effort is not affected. Since we focus on the effects of persistent and transient components of efficiency, we avoid further mathematical complications by assuming that marginal costs of electricity distribution are negligible. Clearly, marginal costs of quality and efficiency effort are still present in the following analysis.

[12]Note that we rule out any firm's strategic behaviour that raises prices to anticipate future constraints and influence efficiency targets. The model assumes that firms do not have information on the timing of introduction of the new price cap regulation. This is equivalent to assume that the regulator do not use the most recent information to set efficiency targets or the discount factor on future profits is relatively large.

From comparison with the unregulated solution in Eq. (6) we see that price cap regulation reduces current price, as expected, because efficiency gains are internalized. In addition, price cap regulation decreases quality since the marginal reward of quality increases is lower. Generally, this motivates the introduction of additional regulatory mechanisms, such as minimum quality standards, that will be later addressed in our analysis. Finally, note that efficiency effort is not affected by the introduction of price cap regulation. This is because we assumed separability between quality and efficiency effort in the cost function to simplify the analysis. Moreover, firms do not know the relationship between the efficiency target (X) set by the regulator and their own efficiency effort (e_t^*). Therefore, efficiency targets are taken as exogenous. Again, this simplifies the analysis and allows us to focus on the possible implications of transient and persistent inefficiency under the price cap regime.

The regulator can set efficiency targets according to two main approaches. The first approach estimates the efficiency target based on the efficiency effort observed in some past period $t-l$: e_{t-l}^*. Since the efficiency effort is the same in all periods before the introduction of the price cap (see Eq. (6)), the authority expects transient efficiency gains of $\hat{E}_T = (1-\alpha)\, e_{t-l}^* = (1-\alpha)\, e_t^*$. In addition, because of persistent inefficiency, part of the current efficiency target can only be achieved in period $t+1+j$, when all the benefits generated by the efficiency effort in period $t+1$ will be exploited. Therefore, in period $t+1$ the authority expects persistent efficiency gains of $\hat{E}_P = \alpha e_{t-l-j}^* = \alpha e_t^*$.

As an alternative, in the second approach to efficiency targets, the regulator can set the efficiency target based on the average efficiency effort observed in other firms, $\bar{e}_t$. Given that firms are all identical, the authority expects transient efficiency gains of $\hat{E}_T = (1-\alpha)\, \bar{e}_t = (1-\alpha)\, e_t^*$. The expected persistent efficiency gains are $\hat{E}_P = \alpha \bar{e}_{t-j}^* = \alpha e_t^*$. In conclusion, both approaches (as well as their combination) generate the same efficiency target. Using Eq. (6), we can write the optimal one-period efficiency target per unit of output as:[13]

$$X^* = \frac{\left(\hat{E}_T + \hat{E}_P\right)}{s(p_t^*, q_t^*)} = \frac{e_t^*}{s(p_t^*, q_t^*)} = \frac{8\left[1-\alpha\left(1-\delta^j\right)\right]}{\theta^3}. \quad (11)$$

[13]This is optimal in the sense that the regulatory mechanism perfectly incorporates the level of efficiency that firms would be able to achieve to maximize their profits in the absense of regulation.

Proposition 2 *The optimal efficiency target decreases with persistent efficiency (α) if the discount factor (δ) on future earnings is less than 1.*

Both mechanisms to set efficiency targets are designed in accordance with attainable cost reductions. The rent of local monopolies in the market for electricity distribution can then be fully extracted to improve consumer welfare. Transient and persistent efficiency gains are fully internalized. The regulator is aware that part of the efficiency target can only be achieved in period $t+1+j$, when all the benefits generated by the efficiency effort in period $t+1$ can be exploited. Therefore, the regulator can measure the persistent component correctly and estimate total efficiency targets achievable in each period. However, this result may be undermined when information on the effects of efficiency effort is incomplete and different types of inefficiency are unobservable.

2.1.2 Imperfect information

If the regulator ignores persistent inefficiency, the estimates of the achievable efficiency target in a price cap environment can be wrong. Consequently, this may lead to quality distortion. Let us assume that the regulation authority is completely ignorant about persistency, i.e. the estimated level of α is $\hat{\alpha} = 0$. Since the delayed effects of today's efficiency effort are neglected, the efficiency target is obtained by substituting α=0 into Eq. (11) as follows:

$$X \mid_{\hat{\alpha}=0} = \frac{\hat{E}_T \mid_{\hat{\alpha}=0}}{s(p_t^*, q_t^*)} = \frac{e_t^* \mid_{\hat{\alpha}=0}}{s(p_t^*, q_t^*)} > X^*. \tag{12}$$

An important implication of imperfect information on efficiency structure is that service quality may be undermined. From Eq. (6) we can see that

$$\partial q_{t+1}^* / \partial X < 0. \tag{13}$$

Since efficiency gains are overestimated, the regulator imposes a tighter price cap. This decreases the marginal benefits of quality and, therefore, the equilibrium level of quality will be lower. Consequently, tighter quality controls are required to avoid too low quality levels.

Whenever some degree of persistent efficiency is present, the regulator cannot set optimal efficiency targets. Even when the regulator is not completely ignorant about the persistent component, efficiency targets will have perverse incentives.

The underestimation of persistency ($\hat{\alpha} < \alpha$) will generally be detrimental for service quality:

$$X \mid_{\hat{\alpha}<\alpha} = \frac{\hat{E}_T + \hat{E}_P \mid_{\hat{\alpha}<\alpha}}{s(p_t^*, q_t^*)} = \frac{e_t^* \mid_{\hat{\alpha}<\alpha}}{s(p_t^*, q_t^*)} > X^* \quad \Rightarrow \quad q_{t+1}^* \left(\hat{X}\right) < q_{t+1}^* \left(X^*\right). \quad (14)$$

Since the price-cost margin is also reduced and profitability is more likely under pressure, firms may postpone important expenditures that may result in persisting poor quality service. Conversely, the overestimation of persistency ($\hat{\alpha} > \alpha$) will give more rent to the monopolistic firm:

$$X \mid_{\hat{\alpha}>\alpha} = \frac{\hat{E}_T + \hat{E}_P \mid_{\hat{\alpha}>\alpha}}{s(p_t^*, q_t^*)} = \frac{e_t^* \mid_{\hat{\alpha}>\alpha}}{s(p_t^*, q_t^*)} < X^* \quad \Rightarrow \quad \pi_{t+1}^* \left(\hat{X}\right) > \pi_{t+1}^* \left(X^*\right). \quad (15)$$

Although quality is higher, overall consumer benefits are lower since the regulator is not successful in extracting the rent from electricity distribution companies.

Proposition 3 *It is impossible to achieve optimal efficiency targets if the regulator is imperfectly informed on persistent efficiency. Underestimation of persistent efficiency leads firms to postpone expenditures, which worsens service quality. Overestimation of persistent efficiency will increase monopoly rents.*

2.1.3 Quality standards

Usually, under price cap regulation and unobservable quality aspects the regulator may consider an additional instrument to regulate firms: Minimum Quality Standards (MQS). These standards can be implemented by periodical controls on service quality to avoid that the price cap mechanism leads to lower quality, i.e. below the socially optimal level of quality or some other satisfactory level of quality.

The combination of price cap and MQS mechanism applies to the case of electricity distribution companies in New Zealand (Shen and Yang, 2012; New Zealand Commerce Commission, 2015). Firms are subject to regulation under the Commerce Act 1986, which defines a price and quality threshold regime since 2001. This regulatory mechanism identifies companies whose performance may warrant further examination. Quality thresholds are based on two criteria: reliability and engagement with consumers to determine their demand for service quality. The reliability criterion requires that unplanned interruptions should not exceed the previous five-year averages. The interruption indicators used are

SAIDI (System Average Interruption Duration Index - minutes per connected customer) and SAIFI (System Average Interruption Frequency Index - interruptions per connected customer). Since 2010, a more rigorous system is in place based on Default Price-quality Path (DPP) and Customized Price-quality path (CPP).

The socially optimal level of quality is used as a reference for instance by the regulator in Norway. In the Norwegian regulation scheme, the network companies' revenue caps are adjusted by quality, i.e. in accordance with the customers' interruption costs (Kjølle et al., 2008). From Eq. (3), we can obtain a welfare function by adding the consumer surplus, $CS_t = \sum_{t=k}^{\infty} \delta^{t-k} \frac{1}{2}(\theta - p_t)s(p_t, q_t)$.[14] For $\gamma = 0$, welfare is maximized at $q_t^{**} = \frac{1}{2}\theta^2 > q_t^* = \frac{1}{4}\theta^2$. To show the possible implications of persistent efficiency under MQS mechanisms, let us assume that service quality can be affected by exogenous shocks (interruptions). Therefore, the true level of quality is:

$$q_{t+1}^* = \hat{q}_{t+1} + \tilde{\epsilon}, \tag{16}$$

where $\hat{q}_{t+1}$ is the observed level of service quality and $\tilde{\epsilon}$ is the exogenous shock distributed as:

$$\tilde{\epsilon} \in \{\epsilon, -\epsilon\}, \quad \Pr[\tilde{\epsilon} = \epsilon] = \rho, \quad \Pr[\tilde{\epsilon} = -\epsilon] = 1 - \rho, \tag{17}$$

where the probability ρ is unknown to the regulator.

Because of exogenous interruptions, the regulator can only verify if the observed level of quality is in the "optimal" range $q_t^{**} - \epsilon \leq \hat{q}_{t+1} \leq q_t^{**} + \epsilon$. Since an observed level of quality below q_t^* is tolerated if it is not too low, i.e. $\hat{q}_{t+1} \geq q_t^{**} - \epsilon = q_{min}$, this provides room for speculation by electricity distribution companies. Under pressure, companies may be more prone to rely on positive shocks, thus speculating on the expected level of service quality. Indeed, the expected level of observable quality is:

$$\hat{q}_{t+1} = \rho(q_{t+1}^* + \epsilon) + (1 - \rho)(q_{t+1}^* - \epsilon) = q_{t+1}^* + \epsilon(2\rho - 1). \tag{18}$$

This may fall below q_{min} if:

$$q_{t+1}^* + \epsilon(2\rho - 1) < q_t^{**} - \epsilon, \tag{19}$$

[14] We assume a fixed cost of regulation enforcement.

which depends on the efficiency target X. Indeed, using Eq. (6) and the socially optimal level of quality $q_t^{**} = \theta^2/2$, we can write the inequality (19) as:

$$\frac{\theta^2}{4}\left[1 - \frac{1}{2}(1-X)^2\right] > \rho\epsilon, \tag{20}$$

which is more likely to be satisfied (i.e. quality below the minimum) for higher values of the efficiency target (X), smaller shocks (ϵ) and lower probability (ρ) of positive shocks (i.e. higher probability of interruptions). Remember that the presence of persistent efficiency reduces the optimal efficiency target (see Proposition 2) Therefore, when the regulator ignores or underestimates persistent efficiency, wrong (too high) efficiency targets may increase violations of MQS, *ceteris paribus*, leading to delayed expenditures and poorer quality in the future.

Proposition 4 *If persistent efficiency is underestimated or completely ignored and service quality is imperfectly observed, firms are less likely to comply with Minimum Quality Standards.*

3 Empirical exercise: cost model specification and estimation method

Our theoretical model hypothesizes that there are two components of inefficiency. If the authority neglects these components, the effectiveness of the regulation may be undermined. Consequently, the effort to separate transient and persistent inefficiency could improve the performance of the electricity distribution market. In the following empirical analysis, we show that data availability and the application of econometric models allow to disentangle transient and persistent inefficiency. Moreover, we provide some evidence that the presence of persistent efficiency may result in poorer quality levels.

The total cost of an electricity distribution company can be specified as a function of input prices and outputs. Moreover, as discussed in Filippini and Wetzel (2014) and in previous studies on the cost structure of electricity distribution companies, in the cost model specification it is important to include a number of output characteristic variables.[15] These variables should take into account the heterogeneity of the electricity distribution companies' production environment.

[15]For a discussion on the estimation of cost functions in the energy sector see Farsi and Filippini (2009).

Generally, the explanatory variables considered in the specification of a cost function for electricity distribution companies are: the quantity of electricity distributed, the number of customers and the factor prices, some output characteristics such as customer density, network size, service area, service quality and load factor.

In this analysis, we specify a total cost function with two outputs and three output characteristics. Unfortunately, the cost model specification does not include input prices since these data are lacking. Consequently, we hypothesize that all electricity distribution companies are exposed to the same input prices.[16]

The total cost can be written as:

$$TC = c\left(Y, CU, NL, LF, Q, T\right), \tag{21}$$

where Y and CU represent the output measured by the electricity supplied in kilowatt-hours and the number of final consumers, respectively. NL, LF and Q are output characteristics: NL is the network length, LF denotes the load factor, and Q is service quality measured by $SAIDI$, an index of the average interruption duration of the system. Finally, T is a time trend that captures changes in the cost over time. In order to be able to compute three type of economies, i.e. economies of output density, economies of customers' density and economies of scale, we use the network length instead of customer density previously used by Filippini and Wetzel (2014). As indicated by the microeconomic theory of production, the cost function should be concave in input prices, non-decreasing in input prices and output, and linearly homogeneous in input prices.

For the estimation of the cost function (21), we use a translog functional form. The translog has the advantage that it does not impose a priori restrictions on the nature of the technology. However, in case the model specification includes some variables relatively highly correlated, then the estimation of the translog cost function can suffer from multicollinearity. In our case, some of the explanatory variables, such as the number of customers, the network length and the load factor, are highly correlated and cause problems in the econometric estimation. Therefore, we estimate a reduced version of the translog, where all interaction

[16]Clearly, we are aware that this is an important assumption. Note, however, that the market for inputs in New Zealand is quite competitive and prices are expected to be similar across distribution companies. The assumption is also used in previous studies as well as by the regulator.

variables have been dropped. Based on Eq. (21) the reduced translog cost function can be expressed as:

$$\begin{aligned}\ln TC_{it} = {} & \beta_0 + \beta_Y \ln Y_{it} + \beta_{CU} \ln CU_{it} + \beta_{LF} \ln LF_{it} + \beta_Q \ln Q_{it} + \\ & + \beta_{NL} \ln NL_{it} + \frac{1}{2}\beta_{YY}(\ln Y_{it})^2 + \frac{1}{2}\beta_{CUCU}(\ln CU_{it})^2 + \\ & + \frac{1}{2}\beta_{LFLF}(\ln LF_{it})^2 + \frac{1}{2}\beta_{QQ}(\ln Q_{it})^2 + \frac{1}{2}\beta_{NLNL}(\ln NL_{it})^2 + \\ & + \beta_t T_t + \varepsilon_{it}, \end{aligned} \tag{22}$$

where the subscripts i and t denote the firm and year, respectively; and the βs are unknown parameters to be estimated. The error term in Eq. (22) is still general and will be specified later from an econometric point of view.

As discussed in more details in Filippini and Greene (2015), several different panel data stochastic frontier models (SFA) can be used to estimate the level of productive efficiency. Some of these models estimate the persistent part of the level of productive efficiency. Others estimate the transient component. Moreover, some recent developed models provide information on whether a firm is characterized by the presence of both types of productive inefficiency.

In this paper, we decided to use three alternative stochastic frontier models: two classical models and one relatively new model. The first model is the basic version of the random effects model proposed by Pitt and Lee (1981) (RE hereafter); the second model is the so-called true random effects model (TRE hereafter) proposed by Greene (2005a, 2005b); and the third model is the generalized true random effects model (GTRE) recently introduced by Colombi et. al. (2014) and Filippini and Greene (2015).

The random effects model introduced by Pitt and Lee (1981) interprets the individual random effects as inefficiency rather than unobserved heterogeneity as in the traditional literature on panel data models. This model provides information on the persistent part of the inefficiency in the use of energy. One problem with the RE is that any unobserved, time-invariant, group-specific unobserved heterogeneity is considered as inefficiency. As a result, this model tends to underestimate the level of persistent efficiency in the use of energy.

The TRE proposed by Greene (2005a and 2005b) extends the SFA model in its original form (Aigner, et al., 1977) by adding an individual random effect in the model. In general terms, for the TRE the constant term, β_0, in Eq. (22), is replaced with a series of firm-specific random effects. This model has the advantage that controls for unobserved heterogeneity that is constant over time. However, any time-invariant component of inefficiency is completely absorbed in the firm-specific constant terms. Therefore, the TRE tends to underestimate the level of inefficiency. Generally, the TRE provide information on the time-varying part of the inefficiency.

The third model (GTRE) offers the possibility to estimate at the same time the persistent and transient part of inefficiency. Colombi et al. (2014) have provided a theoretical platform on which to distinguish persistent from transient inefficiency. Filippini and Greene (2015) suggest a practical completion to the development by proposing a straightforward, transparent empirical estimation method of the GTRE.

Table 1 summarizes the three econometric specifications used in this empirical part of the paper.

4 Data

The data set used in this study is a panel of 28 New Zealand's electricity distribution businesses (EDBs) between 1996 and 2011.[17] The panel is constructed mainly by exploiting information in the "NZ EDB Database" from Economic Insights (Economic Insights, 2009). This database consists of financial and production data on electricity distribution companies. As required by the New Zealand electricity regulation, financial and production data are yearly published in the Electricity Information Disclosures.

In terms of the number of connected customers, the size of companies in our sample varies between 4,100 and 680,000. Total cost is defined as the sum of variable cost and capital cost. Regrettably, consistent information on capital cost for the whole period under observation are not present in the Electricity Information Disclosures. Therefore, we use the so-called optimized deprival value (ODV), which is the annual monetary value of the system fixed assets. The ODV captures the loss of value that a company would bear if deprived of assets.[18] To measure the cost of capital, we follow Lawrence (2003) and Lawrence et al. (2009) and set a common depreciation rate of 4.5 percent of ODV and an opportunity

[17] Few companies have been excluded because of lack of information. Further, a new company recently established has been excluded because of too few years of operation.

[18] The ODV methodology used for asset valuations in the New Zealand's electricity distribution sector is described in detail by the New Zealand Commerce Commission (2004).

cost rate of 8 percent of ODV. Consequently, capital cost is nearly 12.5 percent of the annual ODV.[19] Total cost is adjusted for inflation using the consumer price index for New Zealand provided by the OECD (base 2005).

In addition to the input and output variables, we consider three network characteristics: the load factor, network quality and network length. The load factor captures the intensity in utilization of the distribution network. This is measured by the ratio between the electricity supplied and the maximum distribution transformer demand multiplied by the total number of hours in one year. Lower costs are expected for distribution companies with higher rates of network utilization. Therefore, the estimated coefficient of the load factor is expected to show a negative sign.

The network quality characteristic is measured by SAIDI. This is the average number of interruption minutes for a consumer within a given period. The impact of SAIDI on total costs is rather unclear. On the one hand, higher quality, that is a lower SAIDI, may require more investments and hence may induce higher capital costs. The higher quality may also lead to lower operational costs. For the estimation of our models, we decided to use a weighted level of quality instead of the actual level of quality. During the period under study, the regulator in New Zealand set a minimum level of quality using the previous five-year average of SAIDI. By using the regulated level of SAIDI, we can limit the endogeneity problem related to quality and take into account the relatively high volatility of SAIDI.

Finally, the network length is measured in kilometers to approximate the service area size. We expect a positive coefficient, indicating that companies with a larger area size operate at higher costs than companies with smaller area size do. Some descriptive statistics of the variables used in this study are provided in Table 2.

5 Results

The estimation results for the three models are given in Table 3. These results show that the coefficients of output, number of customers and network length are

[19] Clearly, the limitations arising from the use of ODV as a proxy for capital cost are not ignored. Nevertheless, in the absence of a consistent alternative measure this is plausibly the best proxy for capital cost.

positive and significant across all different estimators. In general, the estimated coefficients are relatively similar across the estimators, except for the coefficients of the two outputs and the coefficient of quality.

Since total costs and regressors are in logarithms and normalized, the first order coefficients are interpretable as cost elasticities evaluated at the sample median. All these coefficients have the expected sign and are highly significant. For instance, the output coefficients suggest that the increase in costs due to a one percent increase in the number of Kwh of electricity distributed, keeping all other explanatory variables constant, varies between 0.16 and 0.33 percent. The coefficient of the network length suggests that the increase in costs due to a one percent extension in the network, keeping all other explanatory variables constant, is approximately 0.2 percent. Further, the coefficient of the number of customers suggests that the increase in costs due to a one percent increase in the number of customers, keeping all other explanatory variables constant, varies between 0.36 and 0.52 percent. The coefficient of the time trend is positive and indicates that total costs of electricity companies increased over time. The cost elasticity with respect to the load factor is negative in all specifications of the cost model, indicating that a 1 percent improvement in the load factor reduces costs by approximately 0.1 percent. Finally, the quality index measured by the regulated level of interruptions (SAIDI) has a negative and significant impact on costs, though this impact is quite small. A 1 percent decrease in quality (i.e. higher number of interruptions) decreases costs between 0.01 and 0.03 percent, *ceteris paribus*. This suggests that reducing service quality allows firms to save on costs.

5.1 Persistent and transient efficiency

The firm's inefficiency for the RE and the TRE models are estimated using the conditional mean of the inefficiency term proposed by Jondrow et al. (1982). Following Filippini and Greene (2015) and using a result from Colombi (2010) based on the moment generating function for the closed skew normal distribution, we compute the inefficiency scores.

Table 4 provides descriptive statistics for the cost efficiency estimates for the 28 electricity distribution companies obtained from the econometric estimation of the three models. The estimation results for the new cost frontier model

(GTRE) provide estimates of the persistent (PGTRE) as well as the transient component of cost efficiency (TGTRE). The RE model produces values of the cost efficiency that are time-invariant and, therefore, should reflect the persistent part of the cost efficiency. On the other hand, the TRE model produces values that are time varying and, therefore, should reflect the transient part of the cost efficiency.

The values reported in Table 4 show that the estimated average values of the persistent efficiency varies from 78 percent in the RE model to 88 percent in the GTRE model. The estimated average values of the transient efficiency varies from 94 percent in the TRE model to 88 percent in the GTRE model. The values of the persistent and transient efficiency obtained by the GTRE model compared to the values obtained by the TRE and RE models are significantly different. This implies that the values obtained by the RE and TRE models do not provide precise information on the level of persistent and transient efficiency. Another interesting comparison across the models is the correlation between inefficiency scores. Although there are differences in the levels of inefficiency, small differences are observed in terms of inefficiency ranking. Table 5 provides the correlations between the estimated levels of cost efficiency obtained from the three model specifications. The value of the correlation coefficients for the transient cost efficiency obtained with TRE and GTRE models is relatively high (0.78). However, the correlation between the values of the persistent cost efficiency obtained with RE and GTRE models is relatively low (0.43). This suggests that the result obtained with the RE model is not measuring the persistent efficiency of the firms correctly. As suggested by Greene (2005b), the reason could be that all unobserved time invariant heterogeneity in the RE model is captured by the individual effect, which is also used to compute the level of efficiency.

The evidence on the presence of persistent efficiency casts doubts on the effectiveness of a price cap regulation that does not distinguish the two parts of efficiency. As suggested by the theoretical model in Section 2 (Proposition 3), when persistent efficiency is ignored the regulator may assume suboptimal efficiency targets. Remember that the higher the share of the persistent component of efficiency, αe^{*}, the lower the share of the transient component of efficiency, (1-α)e^{*}. Moreover, errors in setting efficiency targets may increase with the size of the persistent component as with respect to the transient component.

5.2 Efficiency and quality

Theoretically, the regulatory implications of persistent inefficiency could be assessed by comparing the regulated setting with an ideal setting without price cap and quality regulation. Unfortunately, this experimental design cannot be performed with our dataset. Still, some figures are worth discussing and maybe can stimulate opportunities for future research.

Using the estimated levels of transient and persistent efficiency we can calculate the ratio of the persistent component, α, assumed in the theoretical model as:

$$\frac{\hat{E}_P}{\hat{E}_P + \hat{E}_T} = \frac{\alpha e^*}{\alpha e^* + (1-\alpha)e^*} = \alpha. \tag{23}$$

The theoretical model suggests that, if the regulator is fully informed, service quality is expected to increase with the persistent efficiency component because of the relaxation of the efficiency target. In New Zealand, one of the objectives of the regulator is to provide services at a quality that reflects consumer demand, and the use of total factor productivity should allow setting efficiency targets according to this objective (Brown and Moselle, 2008). Within this framework, the regulator does not disentangle the two parts of efficiency, i.e. the persistent and the transient components. Our results show a positive correlation between the estimated level of persistent efficiency (α) and the estimated level of quality measured by the inverse of SAIDI, although this correlation is very small ($\rho = 0.05$). However, the number of firms that do not comply with the regulated quality level is remarkable.[20] Around 37% of firms on average across the whole period are below the regulated quality standard, ranging from a minimum of 15% in 2001 and a maximum of 68% in 2007. Note that, according to Proposition 4, if persistent efficiency is ignored the level of compliance may suffer. More evidence is provided by the estimated level of α that appears to be negatively correlated with quality compliance, though only weakly ($\rho = -0.04$). Therefore, our results may suggest evidence of errors in the evaluation of persistent efficiency and the setting of efficiency targets.

To further investigate this issue, we can build a measure of the information bias in the estimation of persistent efficiency. We compare the estimates of ineffi-

[20] We calculated the level of quality compliance as the ratio (or the difference) between the minimum quality standard (i.e. average interruptions measured by SAIDI in the previous five-year) and current quality (current level of SAIDI).

ciency from a conventional stochastic frontier pooled model (Aigner et al., 1977) with the estimates of the GTRE model. In most countries (e.g. Austria and Germany), national authorities adopt simple frontier models based on cross-section OLS to estimate the level of efficiency of electricity distribution companies. This may lead to biased results since persistent efficiency is not correctly taken into account. Therefore, we can use the results of the GTRE model as a benchmark of correct information on persistent efficiency and compare them with the efficiency results of a pooled model that ignores the magnitude of persistent efficiency. In this way, we obtain a proxy of the information bias in the estimation of persistent efficiency. As expected, we observe that the larger the information bias the lower is quality compliance. Although this correlation is small ($\rho = -0.09$), the result may cast doubts on the ability of the regulator to ensure quality compliance when an information bias is present on the estimation of persistent efficiency.

5.3 Economies of scale

The estimation results reported in Table 3 can also be used to compute the value of the economies of scale. The inclusion of the number of customers and the network length in the cost function allows for the distinction of economies of output density, economies of customer density and economies of scale. Following Roberts (1986) and Filippini (1998) we define economies of output density (EOD) as the proportional increase in total costs brought about by a proportional increase in output, holding all input prices, the load factor, the number of customers and the network length fixed. This is equivalent to the inverse of the elasticity of total costs with respect to output:

$$EOD_{TC} = \frac{1}{\frac{\partial lnTC}{\partial lnY}}. \tag{24}$$

We find economies of output density if EOD is greater than 1 and, accordingly, we identify diseconomies of output density if EOD is below 1. In the case of EOD = 1, no economies or diseconomies of output density exist. Economies of output density exist if the average cost of an electricity distribution utility decreases as the volume of electricity sold to a fixed number of customers in a service area of a given size increases. This measure is relevant to decide whether side-by-side competition or local monopoly are the most efficient form in the electricity distribution industry.

Economies of customer density (ECD) are defined as the proportional increase in total costs brought about by a proportional increase in output and the number of customers, holding all input prices, the load factor and the network length fixed. Therefore, economies of customer density can be defined as

$$ECD_{TC} = \frac{1}{\frac{\partial lnTC}{\partial lnY} + \frac{\partial lnTC}{\partial lnCU}}. \quad (25)$$

There are economies of customer density if ECD is greater than 1 and, accordingly, we identify diseconomies of scale if ECD is below 1. In the case of ECD = 1 no economies or diseconomies of customer density exist. This measure is relevant for analyzing the cost of distributing more electricity to a fixed service area as it becomes more densely populated.

Economies of scale (ES) are defined as the proportional increase in total costs brought about by a proportional increase in output, the number of customers and the size of the service area, holding all input prices and the load factor fixed. Economies of scale (ES) can thus be defined as

$$ES_{TC} = \frac{1}{\frac{\partial lnTC}{\partial lnY} + \frac{\partial lnTC}{\partial lnCU} + \frac{\partial lnTC}{\partial lnNL}}. \quad (26)$$

There are economies of scale if ES is greater than 1 and, accordingly, we identify diseconomies of scale if ES is below 1. In the case of ES = 1 no economies or diseconomies of scale exist. This measure is relevant for analyzing the impact on cost of merging two adjacent electricity distribution utilities.

Table 6 presents the estimates of economies of scale computed for a medium sized firm.[21] We note that the indicators for economies of scale computed, using the results obtained from the variable cost frontier function, are greater than 1, whereas the values of the economies of scale obtained from the total cost frontier function are approximately equal to 1. However, since the total cost frontier function considers all inputs as freely adjustable, this result may be imprecise. For this reason, we tend to support the hypothesis that the electricity distribution sector is characterized by economies of scale.

[21] Equations (24), (25) and (26) have been evaluated at the values for the load factor, SAIDI and consumer density of the median company. For the interpretation of the results, it is important to note that a proportional increase in electricity supplied and number of consumers implies, keeping the value of the consumer density constant, an increase in the network length.

6 Conclusions

The level of productive efficiency of a firm can be split in two parts: a persistent and a transient component. This distinction can be important in the application of incentive-based regulation schemes, such as the price cap that uses inefficiency scores in the definition of prices in water, electricity and telecommunication sectors. If the regulator ignores or underestimates persistent efficiency, efficiency targets can be wrongly set. As a consequence, this may lead to quality distortion and lower quality compliance.

Generally, the empirical literature on efficiency analysis of firms has not paid a lot of attention to the distinction between these two components. Some scholars (Colombi et al., 2014; Kumbhakar and Tsionas, 2012; Kumbhakar et al., 2012; Filippini and Greene, 2015) have recently proposed econometric approaches to provide separate estimates of the two components of efficiency. Some of these approaches are relatively cumbersome or are based on a multistep manipulation of OLS that is not completely satisfactory. In this paper, we apply the estimator proposed by Filippini and Greene (2015) to assess the level of persistent and transient efficiency in the New Zealand electricity distribution sector. The estimator is based on maximum simulated likelihood using all the sample distributional information to obtain the estimates, and is very effective and strikingly simple to apply.

The empirical results show that the transient and the persistent parts of productive efficiency are relatively different in absolute value and differ from productive efficiency measured by previous approaches. From a regulatory point of view, following the theoretical model, the results imply that differentiated measures of efficiency should be used in regulation. We found some weak evidence that higher levels of persistent efficiency are positively correlated with quality levels. However, electricity distribution companies seem to suffer systematically from poor quality compliance. Moreover, quality compliance is weakly decreasing with both persistent efficiency and the information bias of the regulator in the estimation of persistent efficiency. Further research is needed to confirm these preliminary findings.

References

Aigner D, Lovell CAK, Schmidt P (1977) Formulation and estimation of stochastic frontier production function models. Journal of Econometrics 6(1):21-37

Bernstein JI, Sappington DEM (1999) Setting the X factor in price-cap regulation plans. Journal of Regulatory Economics 16:5-25

Brown T, Moselle B (2008) Use of total factor productivity analyses in network regulation case studies of regulatory practice. The Brattle Group Ltd.

Cowan SGB (1997a) Tight average revenue regulation can be worse than no regulation. The Journal of Industrial Economics 45 (1):75-88

Cowan SGB (1997b) Price-cap regulation and inefficiency in relative pricing. Journal Regulatory Economics 12:53-70

Colombi R (2010) A skew normal stochastic frontier model for panel data. Proceedings of the 45th Scientific Meeting of the Italian Statistical Society, University of Padua, June 29-July 1

Colombi R, Kumbhakar SC, Martini G, Vittadini G (2014) Closed-skew normality in stocastic frontiers with individual effects and long/short-run efficiency. Journal of Productivity Analysis 42:123-136

Di Giorgio L, Filippini M, Masiero G (2015) Structural and managerial cost differences in nonprofit nursing homes. Economic Modelling 51:289-298

Economic Insights (2009) Economic Insights NZ EDB Database. http://www.comcom.govt.nz/assets/Imported-from-old-site/industryregulation/Electricity/PriceQualityPaths/ContentFiles/Documents/comcom-economicinsightedbdatabaseandanalysisdatafiles-aug2009.zip, accessed February 2012

Farsi M, Filippini M (2009) An analysis of cost efficiency in Swiss multi-utilities. Energy Economics 31(2):306-315

Filippini M, Greene W (2015) Persistent and transient productive inefficiency: a maximum simulated likelihood approach. Journal of Productivity Analysis 1-10, DOI 10.1007-/s11123-015-0446-y1-10.

Filippini M, Wetzel H (2014) The impact of ownership unbundling on cost efficiency: Empirical evidence from the New Zealand electricity distribution sector. Energy Economics 45:412-418

Greene W (2005a) Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. Journal of Econometrics 126(2):269-303

Greene W (2005b) Fixed and random effects in stochastic frontier models. Journal of Productivity Analysis 23(1):7-32

Jondrow J, Knox Lovell CA, Materov IS, Schmidt P (1982) On the estimation of technical inefficiency in the stochastic frontier production function model. Journal of Econometrics 19(2-3):233-238

Joskow PL, Schmalensee R (1986) Incentive regulation for electric utilities. Yale Journal on Regulation 4:1

Kjølle GH, Samdal K, Singh B, Kvitastein OA (2008) Customer costs related to interruptions and voltage problems: Methodology and results. IEEE Transactions on Power Systems, 23(3):1030-1038

Kumbhakar SC, Lien G, Hardaker JB (2012) Technical efficiency in competing panel data models: a study of Norwegian grain farming. Journal of Productivity Analysis 41(2):321-337

Kumbhakar SC, Tsionas EG (2012) Firm heterogeneity, persistent and transient technical inefficiency: A generalized true random effects model. Journal of Applied Econometrics 29(1):110-132

Iossa E, Stroffolini F (2002) Price cap regulation and information acquisition. International Journal of Industrial Organization 20:1013-1036

Laffont J-J, Tirole J (1993) A theory of incentives in procurement and regulation. MIT press

Lawrence D (2003) Regulation of electricity lines businesses, analysis of lines business performance 19962003. Report prepared by Meyrick and Associates for the New Zealand Commerce Commission

Lawrence D, Diewert E, Fallon J, Kain J (2009) Electricity distribution industry productivity analysis: 19962008. Report prepared by Economic Insights for the New Zealand Commerce Commission

New Zealand Commerce Commission (2004) Handbook for optimised deprival valuation of system fixed assets of electricity lines businesses

New Zealand Commerce Commission (2015) Regulated Industries: Electricity archive. http://www.comcom.govt.nz/regulated-industries/electricity/, accessed September 2015

Nillesen PHL, Pollitt MG (2011) Ownership unbundling in electricity distribution: empirical evidence from New Zealand. Review of Industrial Organization 38:61-93

Pitt MM, Lee L-F (1981) The measurement and sources of technical inefficiency in the Indonesian weaving industry. Journal of Development Economics 9(1):43-64

Ramos-Real FJ (2005) Cost functions and the electric utility industry. A contribution to the debate on deregulation. Energy Policy 33(1):69-87

Roberts MJ (1986) Economies of density and size in the production and delivery of electric power. Land Economics 378-387.

Shen D, Yang Q (2012) Electricity market regulation reform and competition - Case study of the New Zealand electricity market, In: Wu Y, Shi X, Kimura F (eds) Energy market integration in East Asia: Theories, electricity sector and substitutes, ERIA Research Project Report 2011-17, Jakarta, pp. 103-139

Scully GW (1999) Reform and efficiency gains in the New Zealand electrical supply industry. Journal of Productivity Analysis 11(2):133-147

	Model I	*Model II*	*Model III*
	RE (Pitt and Lee)	**TRE**	**GTRE**
Model	$lnTC_{it} = \beta_0 + \beta'\mathbf{x}_{it} + v_{it} + u_i$	$lnTC_{it} = \beta_0 + w_i + \beta'\mathbf{x}_{it} + v_{it} + u_{it}$	$lnTC_{it} = \beta_0 + (w_i - h_i) + \beta'\mathbf{x}_{it} + v_{it} + u_{it}$
Full random error ε_{it}	$\varepsilon_{it} = u_i + v_{it}$ $u_i \sim N^+(0, \sigma_u^2)$ $v_{it} \sim N(0, \sigma_v^2)$	$\varepsilon_{it} = w_i + u_{it} + v_{it}$ $u_{it} \sim N^+(0, \sigma_u^2)$ $v_{it} \sim N(0, \sigma_v^2)$ $w_i \sim N(0, \sigma_w^2)$	$\varepsilon_{it} = w_i + h_i + u_{it} + v_{it}$ $u_{it} \sim N^+(0, \sigma_u^2)$ $h_i \sim N^+(0, \sigma_h^2)$ $v_{it} \sim N(0, \sigma_v^2)$ $w_i \sim N(0, \sigma_w^2)$
Persistent inefficiency estimator	$E(u_i \mid \varepsilon_{i1}, ..., \varepsilon_{iT})$	None	$E(h_i \mid \varepsilon_{it})$
Transient inefficiency estimator	None	$E(u_{it} \mid \varepsilon_{it})$	$E(u_{it} \mid \varepsilon_{it})$

Table 1: Econometric specifications of the stochastic cost frontier.

Variable	Unit of measurement	Mean	Std. Dev.	Min	Max
Total cost (TC)	Thousand 2005$	$38.5 * 10^3$	$64.0 * 10^3$	$3.4 * 10^3$	$378.0 * 10^3$
Electicity supplied (Y)	MWh	$970.5 * 10^6$	$1753.6 * 10^6$	$37.9 * 10^6$	$10700.0 * 10^6$
Consumers (CU)	Number	62775	114387	4108	679612
Load factor (LF)	Percentage	62.9	7.8	30.4	84.7
SAIDI (Q)	Minutes	221.4	192.5	15.0	1918.0
Network length (NL)	Km	5147.5	6007.3	239.0	30035.5

Number of observations: n=305

Table 2: Descriptive statistics.

Variable	RE	TRE	GTRE
β_0	16.359	16.560	16.449
	(499.981)	(1474.362)	(1380.217)
lnY	0.164	0.241	0.330
	(1.602)	(11.657)	(16.345)
$lnCU$	0.516	0.425	0.361
	(4.210)	(18.300)	(16.140)
$lnLF$	-0.139	-0.122	-0.144
	(-1.190)	(-2.785)	(-3.478)
lnQ	-0.011	-0.015	-0.025
	(-0.580)	(2.183)	(3.816)
$lnNL$	0.210	0.207	0.213
	(2.494)	(15.017)	(16.526)
$lnY*lnY$	0.277	0.202	0.043
	(1.147)	(2.155)	(0.460)
$lnCU*lnCU$	0.424	0.326	0.211
	(2.023)	(3.435)	(2.227)
$lnLF*lnLF$	-0.257	-0.164	-0.150
	(-0.499)	(-0.619)	(-0.719)
$lnQ*lnQ$	0.011	0.014	0.033
	(0.335)	(0.934)	(2.421)
$lnNL*lnNL$	-0.143	-0.157	-0.181
	(-1.686)	(-15.810)	(-18.351)
$lnY*lnCU$	-0.333	-0.232	-0.088
	(-1.441)	(-2.548)	(-0.976)
T	0.020	0.020	0.021
	(11.868)	(22.558)	(21.754)
σ_w	-	0.192	0.189
	-	(36.563)	(37.417)
λ	4.981	1.734	4.933
	(2.030)	(6.130)	(6.320)
σ^2	0.328	0.914	0.119
	(4.489)	(20.260)	(21.530)
σ_h	-	-	0.888
	-	-	(13.471)
Log likelihood	336.692	335.893	328.043

Table 3: Estimated first and second order coefficients from cost frontier models (Asymptotic t-ratios in parentheses).

Variable	Mean	Std. Dev.	Minimum	Maximum
RE	0.782	0.143	0.515	0.984
TRE	0.940	0.032	0.803	0.987
TGTRE	0.878	0.062	0.644	0.990
PGTRE	0.884	0.021	0.866	0.946

Table 4: Cost efficiency scores.

	RE	TRE	TGTRE	PGTRE
RE	1	0.031	-0.235	0.425
TRE	0.315	1	0.779	-0.069
TGTRE	-0.235	0.779	1	-0.653
PGTRE	0.425	-0.069	-0.653	1

Table 5: Correlation coefficients.

	Economy of density	Economy of consumer density	Economy of scale
RE	6.089	1.470	1.123
TRE	4.142	1.698	1.127
GTRE	3.031	1.902	1.119

Table 6: Economies of scale.

IdEP Economic Papers

The series **IdEP Economic Papers**, ideally continues the work of the "Quaderni della Facoltà" the publication of which began in 1998 and ended in 2013.
For a complete list of the Quaderni see: http://econpapers.repec.org/paper/lugwpaper/

IdEP Economic Papers also gathers the legacy of the CEPRA Working Papers, published from 2012 to 2013.
The full list at: http://econpapers.repec.org/paper/lugwcepra/

2014:

No. 01
F. Mazzonna, F. Peracchi, *Unhealthy retirement? Evidence of occupation heterogeneity*

No. 02
L. Di Giorgio, M. Filippini, G. Masiero, *The relationship between costs and quality in nonprofit nursing homes*

No. 03
F.C. Billari, V. Galasso, *Fertility decisions and pension reforms : evidence from natural experiments in Italy*

No. 04
M. Jametti, M. Joanis, *Elections and de facto expenditure decentralization in Canada*

No. 05
M. Jametti, *Weathering the global financial crisis : is direct democracy of any help?*

No. 06
U. Pagano, M. Vatiero, *Costly institutions as substitutes : novelty and limits of the Coasian approach*

No. 07
R. Parchet, *Are local tax rates strategic complements or strategic substitutes?*

No. 08
R. Ippoliti, M. Vatiero, *An analysis of how 2002 judicial reorganization has impacted on the performance of the First Instance Courts* (Preture) *in Ticino*

No. 09
F. Mazzonna, P. Salari, *Short term effects of public smoking bans on health*

2015:

No. 01
S. Galletta, *Direct democracy, partial decentralization and voter information : evidence from Swiss municipalities*

No. 02
I. Sarman, *Second homeowners' intention to move : an integrated ordered logit model with latent variable*

No. 03
M.J. Roe, M. Vatiero, *Corporate governance and its political economy*

No. 04
M. Filippini, M. Koller, G. Masiero, *Competitive tendering versus performance-based negotiation in Swiss public transport*

No. 05
O. Giuntella, F. Mazzonna, *If you don't snooze you lose health and gain weight : evidence from a regression discontinuity design*

No. 06
M. Filippini, B. Hirl, G. Masiero, *Rational habits in residential electricity demand*

2016:

No. 01
S. Galletta, *Law enforcement, municipal budgets and spillover effects : evidence from a quasi-experiment in Italy*

No. 02
S. Galletta, A. Redonda, *Corporate flat tax reforms and businesses' location choices. Evidence from Switzerland*

No. 03
M. Filippini, W. Greene, G. Masiero, *Persistent and transient productive inefficiency in a regulated industry : electricity distribution in New Zealand*