An efficiency correction model

Francke, M.K.; de Vos, A.F.

Publication date
2009

Document Version
Submitted manuscript

AN EFFICIENCY CORRECTION MODEL

August 4, 2009

Marc K. Francke and Aart F. de Vos

University of Amsterdam Business School, Faculty of Economics and Business,
Roetersstraat 11, 1018 WB Amsterdam, the Netherlands
m.k.francke@uva.nl

Vrije Universiteit Amsterdam, Department of Econometrics,
De Boelelaan 1105, NL-1081 HV Amsterdam
avos@feweb.vu.nl

AN EFFICIENCY CORRECTION MODEL

Abstract

We analyze a dataset containing costs and outputs of 67 American local exchange carriers in a period of 11 years. This data has been used to judge the efficiency of BT and KPN using static stochastic frontier models. We show that these models are dynamically misspecified. As an alternative we provide an efficiency correction model. This model makes it possible to distinguish between unmeasured firm heterogeneity, firm inefficiency and measurement error, by assuming time invariant unmeasured firm heterogeneity and firm efficiency evolving over time.

Keywords: Error Correction; Panel Data; Stochastic Frontier
JEL Classification: C23, D2

1 INTRODUCTION

In this paper we introduce a stochastic frontier model for panel data that combines unobserved heterogeneity (or specification error) and firm specific stochastic dynamic efficiency changes. The similarity to error correction models led to the name efficiency correction model (EFCOM). The main focus of the model is to identify the relative efficiency of a firm over time.

We apply the model on two sets of panel data that have previously been analyzed by using models without firm specific dynamics to estimate cost efficiencies. Both datasets consist of cost and output variables per year for several firms. The first dataset relates to 67 U.S. local exchange carriers (LEC's) over the years 1996–2006, while the second dataset relates to 382 U.S. nonteaching hospitals over the years 1987–1991. For both datasets we show that the models without firm specific efficiency dynamics are clearly misspecified and that the efficiency correction model provides a much better fit to the data. Moreover there is a striking similarity in outcomes for both datasets.

Following Griffin and Steel (2007) the main estimation results are obtained by the Bayesian package WinBUGS, using Markov Chain Monte Carlo (MCMC) techniques. However we start our analysis with a classical explanation and estimation procedure of the main dynamic features. We show that maximum likelihood estimates and MCMC inference give very similar results in a simple dynamic model. In order to obtain estimates for all stochastic components in the efficiency correction model – the unobserved heterogeneities and efficiencies – MCMC methods are required, and WinBUGS does the task (though possibly not in the most efficient way). We obtain robust estimates of the relative efficiencies per year.

The setup of this article is as follows. In section 2 we discuss the efficiency correction model in relation to the literature on stochastic frontier models. Section 3 provides a brief description of the LEC data. Section 4 provides a preliminary classical analysis and a comparison between maximum likelihood and WinBUGS outcomes for the LEC data from a simple dynamic stochastic frontier model. Sections 5 and 6 provide estimation results from the efficiency correction model for the local exchange carriers and hospitals, respectively. Section 7 concludes.

2 MODEL SPECIFICATIONS

2.1 A general stochastic frontier model for panel data

We consider a general stochastic frontier model for panel data given by

$$y_{it} = \mu + x_{it}'\beta + \gamma_t + u_{it} + v_i + \varepsilon_{it},\ u_{it} \geq 0, \tag{1}$$

for $i = 1, \ldots, N$, $t = 1, \ldots, T$, where N is the number of firms, T is the number of time periods, and x_{it} is a $(k \times 1)$ vector of explanatory variables. u_{it}, v_i and ε_{it} are respectively inefficiency, unobserved heterogeneity (or specification error) and measurement error, and efficiency is defined as $\exp(-u_{it})$. It is assumed that u_{it}, v_i, and ε_{it} are distributed independently of each other, and $\text{Cov}(\varepsilon_{it}, \varepsilon_{js}) = 0$ for $s \neq t$. γ_t is a deterministic function of time or a stochastic process, being independent of u_{it}, v_i, and ε_{it}. In our applications y_{it} will stand for the logarithm of total cost and x_{it} is a vector of output components and environmental variables. The same model can be used for output or profit, in that case $u_{it} \leq 0$.

The main problem of the model is identification of the stochastic components. Questions are whether it makes sense to distinguish between specification and measurement errors, whether γ_t can be attributed to inefficiency, and whether the inefficiency term u_{it} can be distinguished from other terms. In most studies rather restrictive assumptions are used to obtain identification. These restrictions turn out to be too restrictive, at least in our applications.

Specification and measurement errors have the plausible structure $v_i + \varepsilon_{it}$. The fact that we call v_i the specification error is partly arbitrary; we could also include a part of ε_{it} in the specification error. For the identification of efficiency u_{it} – the goal of the analysis – this is of little importance. What is important is that we assume that $v_i + \varepsilon_{it}$ has no autocorrelation structure. In our applications ε_{it} will appear to be small, supporting the idea that they are indeed measurement errors.

The firm independent time term, γ_t, is well identified in our applications with many firms and short time periods in the form of time dummies, so implicitly in other specifications. Whether these time effects must be attributed to general efficiency changes or other shifts in the cost function cannot be decided on statistical grounds. One may say that absolute inefficiency is unidentified. Relative inefficiency u_{it} per period however, is identified. So that is the output we concentrate on.

Remains the distinction between v_i (or $v_i + \varepsilon_{it}$) and u_{it}. There are two statistical sources of information to identify v_i and u_{it}. The first is that u_{it} is positive and skewed to the right, while the other terms are symmetrical. This however is notoriously weak information. The other possibility is to assume autocorrelation for the inefficiencies. Like Ahn, Good, and Sickles (2000) we will argue that autocorrelation in u_{it} is plausible. These authors did however not include unobserved heterogeneity in their model specification. And we think that, as it is impossible to specify a unique and complete model for cost components, unobserved heterogeneity v_i should be included.

A detailed discussion on unmeasured heterogeneity in stochastic frontier models is provided by Greene (2005). Farsi, Filippini, and Greene (2006) show in a static context the impact of the inclusion of unobserved heterogeneity v_i on the estimates of efficiency. They show that models including v_i underestimate efficiency and models without v_i overestimate efficiency, providing lower and upper bounds for companies' efficiency scores. In a specific example they show that the correlation between the efficiencies scores and ranks in both models is approximately zero.

We will confine ourselves to model (1). We do not consider the case where x_{it} is stochastic and β is allowed to vary over time and clusters, see for example Tsionas (2002) and Kumbhakar and Tsionas (2005), who distinguish between technical and allocative efficiency, allowing β to vary over clusters, and Tsionas and Kumbhakar (2004) where β is allowed to vary over time and clusters in a Markov switching stochastic frontier model. Neither will we investigate dynamics in the form $u_{it} = u_i\omega_t$, so assuming a common (deterministic) function for the evolution of inefficiency over time, see for instance Cornwell, Schmidt, and Sickles (1990), Battesse and Coelli (1992), and Griffin and Steel (2007).

Examples of models including stochastic dynamics (but without unobserved heterogeneity) in the literature are Ahn, Good, and Sickles (2000), using the generalized method of moments for estimation, Desli, Ray, and Kumbhakar, using maximum likelihood estimation, Tsionas (2006), using Gibbs sampling for inference, and Park, Sickles, and Simar (2003) and Park, Sickles, and Simar (2007), using nonparametric estimation methods.

We include unobserved heterogeneity. Following Farsi, Filippini, and Greene (2006) we assume that the specification error is time independent and inefficiency is time dependent, but realize that this assumption is hardly testable.

2.2 Autocorrelation and efficiency correction

The general stochastic frontier model (1) can equivalently be represented in levels for the first period, and first differences for the other periods, leading to

$$y^*_{i1} = \mu + x'_{i1}\beta + \gamma_1 + v_i + u_{i1}, \tag{2}$$

$$\Delta y^*_{it} = \Delta(x'_{it}\beta + \gamma_t) + \Delta u_{it},\ t = 2, \ldots, T, \tag{3}$$

where $y^*_{it} = y_{it} - \varepsilon_{it}$, and Δ is defined as the first difference operator, $\Delta z_t = z_t - z_{t-1}$.

The assumption that $\text{Cov}(u_{it}, u_{is}) = 0$ for $s \neq t$ is very unlikely to hold. Firms do not adjust their cost immediately to output. Lag structures are needed. This idea may be formalized in a model having an error correction structure:

$$y^*_{i1} = \text{EF}_{i1} + u_{i1}, \tag{4}$$

$$\Delta y^*_{it} = -\delta_1(y^*_{i,t-1} - \text{EF}_{it}) + \delta_2\Delta\text{EF}_{it} + \eta_{it},\ \eta_{it} \geq 0,\ t = 2, \ldots, T, \tag{5}$$

$$\text{EF}_{it} = \mu + x'_{it}\beta + \gamma_t + v_i, \tag{6}$$

where it is assumed that $0 \leq \delta_1, \delta_2 \leq 1$, and EF_{it} is the efficiency frontier. The term $u_{i,t-1} = y^*_{i,t-1} - \text{EF}_{i,t-1}$ provides the deviation from the efficiency frontier in the previous period. This is similar to the error correction specification, where it is assumed that u_{i1} and η_{it} are normally distributed and the error correction term is an adjustment to the equilibrium. The second term $\Delta\text{EF}_{it} = \Delta(x'_{it}\beta + \gamma_t)$ is the change in the efficiency frontier. In the error correction model this would be the change in equilibrium. In (4)–(6) firms are allowed to partially adjust their cost to changes in the efficiency frontier. Note that for $\delta_1 = 1$ firms adjust their cost immediately to deviations from the efficiency frontier in the previous period, and for $\delta_2 = 1$ firms fully adjust their cost to changes in the efficiency frontier.

Eq. (4)–(6) can be formulated in unobserved component format as

$$\begin{aligned} y_{it} &= \mathrm{EF}_{it} + u_{it} + \varepsilon_{it}, && (7) \\ u_{it} &= (1-\delta_1)u_{i,t-1} - (1-\delta_2)\Delta\mathrm{EF}_{it} + \eta_{it}, && (8) \end{aligned}$$

showing the unfortunate consequence that, unless $\delta_2 = 1$, inefficiency u_{it} is no longer restricted to be positive.

In order to retain the preferred AR(1) structure for the inefficiencies it is necessary to redefine the efficiency frontier. The model that is given by

$$\begin{aligned} y_{it} &= \mathrm{EF}^*_{it} + u^*_{it} + \varepsilon_{it}, && (9) \\ u^*_{it} &= (1-\delta_1)u^*_{i,t-1} + \eta_{it}, && (10) \end{aligned}$$

is an equivalent representation of (7) and (8) if

$$\mathrm{EF}^*_{it} = \mu + x'_{it}\beta + \gamma_t - (1-\delta_2)\sum_{j=0}^{t-2}(1-\delta_1)^j\Delta(x'_{i,t-j}\beta - \gamma_{t-j}) + v_i.$$

So the efficiency frontier is adjusted for the reaction on recent changes in the equilibrium. In order to stay close to the usual definition of efficiency it seems wise to restrict this adjustment to one period. In that case one obtains

$$\begin{aligned} y_{it} &= \mathrm{EF}_{it} + u_{it} + \varepsilon_{it},\ t = 2,\ldots,T, && (11) \\ \mathrm{EF}_{it} &= \mu + x'_{it}\beta + \gamma_t - (1-\delta_2)\Delta x'_{it}\beta + v_i, && (12) \\ u_{it} &= (1-\delta_1)u_{i,t-1} + \eta_{it},\ \eta_{it} \geq 0, && (13) \end{aligned}$$

which we call the efficiency correction model (EFCOM). Note that we dropped for convenience the asterisks in the notation for the efficiency frontier and inefficiency. We also redefined γ_t. It now stands for the yearly shift in the frontier; to include a partial adjustment to the shift in the past year leads to ambiguity and bad identification.

The efficiency correction model can be seen as a generalization of the model that was proposed by Ahn, Good, and Sickles (2000), having no unobserved heterogeneity v_i in their model specification and assuming that firms adjust their cost immediately to changes in the efficiency frontier ($\delta_2 = 1$).

Additional assumptions has to be made on the initial inefficiencies and the distributions of the innovations η_{it}. We assume covariance stationarity, implying that the unconditional moments are provided by

$$\begin{aligned} \mathrm{E}(u_{it}) &= \mathrm{E}(\eta_{it})/\delta_1, & (14)\\ \mathrm{Var}(u_{it}) &= \mathrm{Var}(\eta_{it})/(1-(1-\delta_1)^2). & (15)\end{aligned}$$

In general the unconditional distribution is not the same as the distribution of the innovations. In the case that η_{it} has a gamma distribution, denoted by $\eta_{it} \sim \mathrm{G}(\phi, \lambda)$, where $\mathrm{E}(\eta_{it}) = \phi/\lambda$ and $\mathrm{Var}(\eta_{it}) = \phi/\lambda^2$, the unconditional distribution can be approximated by a gamma distribution, $u_{it} \sim \mathrm{G}(\lambda(2-\delta_1), \phi(2-\delta_1)/\delta_1)$, following from the moment conditions (14)–(15), which can be checked by simulation. In our applications we assume that the innovations η_{it} have a gamma distribution.

An alternative to the first order autoregressive process [AR(1)] with gamma innovations (13) is provided by Tsionas (2006), assuming that the log of inefficiency follows a first order autoregressive process, i.e. $\ln u_{it} = z_{it}'\gamma + (1-\delta)\ln u_{i,t-1} + \eta_{it}$, where $\eta_{it} \sim N(0, \sigma_\eta^2)$ and z_{it} is a vector of additional explanatory variables.

Another option for the AR(1)-process would be to use a conditional Gamma model for the inefficiency process, replacing Eq. (13) by $u_{it}|u_{i,t-1} \sim G(\phi, \phi/m(u_{i,t-1}))$, where $m(u_{i,t-1}) = \mathrm{E}(u_{it}|u_{i,t-1}) = (1-\delta_1)u_{i,t-1} + \mu_i\delta_1$ and μ_i is the unconditional expectation of u_{it}. The unconditional variance is provided by

$$\mathrm{Var}(u_{it}) = \frac{\mu^2/\phi}{1-(1-\delta_1)^2(\phi+1)/\phi}, \text{ for } 0 \leq 1-\delta_1 < \left(\frac{\phi}{\phi+1}\right)^{1/2} < 1,$$

see for more details Grunwald et al. (2000).

3 DATA

Stochastic frontier models have been used in practice to compare cost efficiency for fixed line telecommunication operators. Two recent examples of these benchmark studies are performed by NERA (2005, 2006) under the authority of Ofcom (Office of Communication) and OPTA (Independent Post and Telecommunications Authority) to judge the cost efficiency of the British BT and the Dutch KPN respectively. These studies are based on the costs and outputs of American local exchange carriers (LEC), which data is freely available.

In this article we use the data from 67 LEC's over a period of 11 years, from 1996 until 2006. The costs are the sum of operating costs, depreciation and cost of capital. The output is measured by the number of switched and leased lines, switch minutes, and the length of cable sheath. Environmental explanatory variables are the proportion of business to residential lines and the population density. All variables are measured in natural logs. The variables and their abbreviations are given in table 1. Detailed information on the different variables can be found in the report by NERA (2006). Table 2 contains averages of the variables over the LEC's per year. The averages of cost, leased lines, sheath, business to residential ratio and population density increase over time, while the averages of switched lines and switch minutes decrease over time. There are large differences between LEC's with respect to cost and output variables; the difference between minimum and maximum cost is a factor 15. Some values of switch minutes and depreciation cost are missing for some LEC's in the years 2005 and 2006. The missing values are replaced by estimates, based on an interpolation or extrapolation of the specific series.

In our applications the leased lines, switch minutes and sheath are specified in deviation from switched lines, indicated by by asterisks, in which case the coefficient of switched lines refers to the economies of scale.

4 PRELIMINARY MODEL EXPLORATION

The purpose of this section is threefold. First it is shown that a simple informal ordinary least squares (OLS) test reveals that a random effects model (REM) without autocorrelation is dynamically misspecified. Next an error correction random effects model (ECREM) is introduced, reflecting the basic dynamic structure of the efficiency correction model (11)–(13). Finally it is shown that for the (error correction) random effects models there is a close correspondence between the estimation results obtained by maximum likelihood and WinBUGS, the program for Bayesian inference that will be used for the efficiency correction model in the next section.

The random effects model without autocorrelation is provided by

$$y_{it} = \mu + x_{it}'\beta + \gamma_t + \theta_i + \alpha_{it},\ \theta_i \sim N(0, \sigma_\theta^2) \text{ and } \alpha_{it} \sim N(0, \sigma_\alpha^2), \tag{16}$$

where it is assumed that θ and α are uncorrelated. We use the symbols θ and α to indicate that in this section no distinction is made between inefficiency, unobserved heterogeneity and measurement error.

The data can be split in two independent parts, in means and deviation from means,

$$\overline{y}_{i.} = \mu^* + \overline{x}_{i.}'\beta + \theta_i + \overline{\alpha}_{i.}, \tag{17}$$

$$\widetilde{y}_{it} = \widetilde{x}_{it}'\beta + \widetilde{\gamma}_t + \widetilde{\alpha}_{it}, \tag{18}$$

where $\overline{z}_{i.} = T^{-1}\sum_{t=1}^{T} z_{it}$, and $\widetilde{z}_{it} = z_{it} - \overline{z}_{i.}$ for $z = y,\ x, \gamma$, and α, and $\mu^* = \mu + \bar{\gamma}_.$. The two sources of information, means and deviations from the means, are independent and should reinforce each other. The assumption that β is the same in (17) and (18) can be informally checked by comparing the OLS estimates of β for both equations. The estimate of β in the random effects model (16) is a weighted average of the estimates of β in (17) and (18), with weights depending on σ_θ^2 and σ_α^2.

Estimates of the variances σ_α^2 and σ_θ^2 may be obtained from regression on the means (17) and deviations (18) from the means: $\hat{\sigma}_\alpha^2 = RSS_D/(N(T-1)-k)$, and $\hat{\sigma}_\theta^2 = RSS_M/(N-k) - \hat{\sigma}_\alpha^2/T$, where RSS_M and RSS_D are the residual sum of squares from (17) and (18), respectively.

The first three panels in table 4 present estimation results for the REM (16) and the model in means (17), and deviations (18) for the LEC data. The models in means and deviations are estimated by OLS, the REM is estimated by maximizing the concentrated (for μ, β, γ, and σ_α) loglikelihood with respect to $q_\theta = \sigma_\theta^2/\sigma_\alpha^2$, indicated by ML in table 4.

The estimation results for β are very different from each other. We focus on the most important variable ln(SL), reflecting the economies of scale. The economies of scale parameter in the means model is, as expected, 0.98, approximately one, while in the deviations model this parameter is only 0.59. The information from the means is dominant: in the REM the estimate of the economies of scale parameter is almost 1.

When σ_θ and σ_α in REM are estimated from the model in means (17) and deviations (18), one obtains $\hat{\sigma}_\alpha = 0.057$ and $\hat{\sigma}_\theta = 0.120$. The estimate $\hat{\sigma}_\theta$ contrasts to the maximum likelihood estimate from REM, being 0.271, another indication that REM is misspecified.

The fourth panel of table 4 presents the Bayesian estimation results for the random effects model (16), obtained by WinBUGS. Noninformative normal distributed priors are assumed for the place parameters μ, β, and γ_t and noninformative gamma distributed priors for σ_θ^{-2} and σ_α^{-2}. The Bayesian estimation results are almost the same as the results from maximum likelihood. WinBUGS also provides the model selection criterion DIC, see Spiegelhalter et al. (2002). DIC is minus two times the loglikelihood in the posterior means plus two times a penalty for model complexity, measured by the "effective number of parameters" and denoted by p_D.

It is clear from table 4 that the random effects model (16) is misspecified. As mentioned in subsection 2.2 this may be the result from the unrealistic assumption in the random effects model that firms immediately adjust their cost to output.

This assumption is relaxed in the error correction random effects model (ECREM), provided by

$$y_{i1} = \mu + x_{i1}'\beta + \gamma_1 + \theta_i + \alpha_{i1},\ \theta_i \sim N(0, \sigma_\theta^2), \tag{19}$$

$$\Delta y_{it} = -\delta_1(y_{it} - \mu - x_{it}'\beta - \gamma_t - \theta_i) + \delta_2\Delta(x_{it}'\beta + \gamma_t) + \eta_{it},\ \eta_{it} \sim N(0, \sigma_\eta^2), \tag{20}$$

where θ and η are uncorrelated. Further we assume covariance stationarity for α_{it}, so

$$\alpha_{i1} \sim N(0, \sigma_\eta^2/(1-(1-\delta_1)^2)). \tag{21}$$

For $\delta_1 = \delta_2 = 1$ the error correction random effects model (19)–(21) coincides with the random effects model (16).

The first three panels of table 5 contain the estimation results from ECREM for three different cases, namely $\delta_2 = 1$, $\delta_1 = \delta_2$, and the general case $0 < \delta_1, \delta_2 < 1$. The models are estimated by maximizing the concentrated (for μ, β, γ and σ_η) loglikelihood with respect to $q_\theta = \sigma_\theta^2/\sigma_\eta^2$, δ_1 and δ_2. By introducing only 2 variables the loglikelihood increases with more than 228 points compared to the random effects model (16). The difference in loglikelihood between the most and least restrictive model is more than 30 points at the cost of only 1 parameter. The estimates of δ_1 in all three cases are far from 1, the value that is assumed in the random effects model. So it can be concluded that the REM is dynamically misspecified. The fourth panel of table 5 contains the Bayesian estimation results for the case $\delta_2 = 1$. Like in the random effects model noninformative priors for place and scale parameters are assumed. For δ_1 a uniform prior between 0 and 1 is assumed. The results almost coincide with the estimation results by maximum likelihood in the first panel. The DIC decreases with 151 points compared to the random effects model, being consistent with the increase in loglikelihood.

5 EFFICIENCY CORRECTION MODEL

In this section the estimation results from the efficiency correction model (11)–(13) are presented. Compared to the error correction random effects model in the previous section the essential new element is the inclusion of measurement errors and the interpretation of the different error components.

We choose to omit the first year in the evaluation of the likelihood because x_{i0} in Δx_{i1} is unknown. An alternative would have been to include the first year and replace $\Delta x'_{i,1}\beta$ by a stochastic variable, with moments determined by those of $\Delta x'_{i,t}\beta$ in later years, so assuming a model for $\Delta x'_{i,t}$. This approach complicates the estimation procedure and has little effect on the results as the first observation gets little weight in the evaluation.

We assume that γ_t in the efficiency frontier (12) follows a random walk plus drift,

$$\gamma_t = \gamma_{t-1} + \kappa + \zeta_t,\ t = 3, \ldots T, \tag{22}$$

where ζ_t are iid with zero mean, and the initial value $\gamma_2 = 0$ as the efficiency frontier already contains a constant term. The random walk with drift specification requires two additional parameters: the drift term κ and the variance σ_ζ^2. One can use dummies as well, but this specification is more parsimonious and elegant as it favors a relatively smooth sequence of γ_t, depending on the variance σ_ζ^2.

We make the following distributional assumptions for the the efficiency correction model, given by Eq. (11)–(13) and (22). The measurement errors ε_{it} have a student distribution to accommodate for possible outliers. The specification errors v_i and the innovations of the random walk ζ_t have a normal distribution and the initial inefficiencies u_{i2} and the innovations η_{it} have a gamma distribution. This can be summarized by

$$\begin{array}{llllllll} \varepsilon_{it} & \sim & t_\nu(0, \sigma_\varepsilon^2), & v_i & \sim & N(0, \sigma_v^2), & \zeta_t & \sim & N(0, \sigma_\zeta^2), \\ u_{i2} & \sim & \text{Gamma}(\phi_1, \lambda_1), & \eta_{it} & \sim & \text{Gamma}(\phi, \lambda). \end{array}$$

From the covariance stationarity assumptions (14)–(15) it follows that

$$\phi = \frac{\delta_1}{2-\delta_1}\phi_1, \text{ and } \lambda = \frac{1}{2-\delta_1}\lambda_1.$$

The parameters to be estimated are $\mu, \beta, \kappa, \delta_1, \delta_2, \phi_1, \lambda_1, \sigma_v, \sigma_\zeta, \sigma_\varepsilon$, and ν. Before specifying priors for these parameters it is useful to examine how well the parameters are identified. As the place parameters are well identified, we compute first and second moments for the efficiency correction model without β, δ_2 and γ_t, given by

$$\begin{aligned} y_{it} &= \mu + v_i + u_{it} + \varepsilon_{it}, \\ u_{it} &= (1-\delta_1)u_{i,t-1} + \eta_{it}, \end{aligned}$$

which may be rewritten as

$$\begin{aligned} y_{i2} &= \mu + v_i + u_{i2} + \varepsilon_{i2}, \\ \Delta y_{it} &= -\delta_1 u_{i,t-1} + \eta_{it} + \varepsilon_{it}\varepsilon_{i,t-1}. \end{aligned}$$

This model contains 7 parameters: $\mu, \delta_1, \phi_1, \lambda_1, \sigma_v, \sigma_\varepsilon$, and ν. The first and second moments of y_{i2} are provided by

$$\begin{aligned} \text{E}(y_{i2}) &= \mu + \phi_1/\lambda_1, \\ \text{Var}(y_{i2}) &= \sigma_v^2 + \phi_1/\lambda_1^2 + \nu/(\nu-2)\sigma_\varepsilon^2. \end{aligned}$$

The other relevant moments follow from the ARMA(1,1) structure of Δy_{it}, and are given by

$$\begin{aligned} \text{Var}(\Delta y_{it}) &= 2(\delta_1\phi_1/\lambda_1^2 + \nu/(\nu-2)\sigma_\varepsilon^2), \\ \text{Cov}(\Delta y_{it}, \Delta y_{i,t-1}) &= -\delta_1^3\phi_1/\lambda_1^2 - \nu/(\nu-2)\sigma_\varepsilon^2, \\ \text{Cov}(\Delta y_{it}, \Delta y_{i,t-2}) &= -\delta_1^3(1-\delta_1)\phi_1/\lambda_1^2. \end{aligned}$$

Note that $\text{E}(\Delta y_{it}) = 0$ due to the stationarity requirement.

The identification of the degrees of freedom ν follows from the fourth moments. In case of outliers, ν will be low. Given ν the parameters $\phi_1/\lambda_1^2, \delta_1, \sigma_\varepsilon^2$ and σ_v can be identified from the variances and covariances equations. The problem lies in the expectation of y_{i2}, the distinction between μ and the mean of the inefficiency ϕ_1/λ_1, see also Griffin and Steel (2007). The essential additional information must come from the third moment of u_{i2}, which has skewness $2/\sqrt{\phi_1}$. This information however is weak, at least in our data. In our applications the MCMC chain does not converge without an additional restriction.

In order to obtain identification of μ and ϕ_1/λ_1 it is required that $P(u_{i2} < 0.05) = 0.02$, so that one expects that 1 in 50 companies has an inefficiency lower than 0.05. A numerical approximation of this restriction is $\lambda_1 = -0.52 - 0.75\phi_1 + 1.58\phi_1^2$. This means that absolute efficiencies cannot be estimated, only relative efficiencies.

For the remaining parameters practically noninformative priors are specified:

$$\begin{array}{llllll} \sigma_\varepsilon^{-2} & \sim & \text{Gamma}(0.001, 0.001), & \nu & \sim & \text{Exp}(1/3), \\ \sigma_\zeta^{-2} & \sim & \text{Gamma}(0.001, 0.001), & \kappa & \sim & N(0, 100), \\ \sigma_v^{-2} & \sim & \text{Gamma}(0.001, 0.001), & \mu & \sim & N(0, 100), \\ \phi_1 & \sim & \text{Uniform}(1, 6), & \beta & \sim & N(0, 100), \\ \delta_1 & \sim & \text{Uniform}(0.5, 0.95), & \delta_2 & \sim & \text{Uniform}(0.5, 0.95). \end{array}$$

The efficiency correction model is estimated by WinBUGS. The inclusion of the measurement errors ε_{it} offers the possibility to generate the specification errors v_i and the efficiencies u_{it} as "parameters" in the MCMC process and to use the distribution of the measurement errors ε_{it} for the computation of the likelihood. Other setups are possible but more complex, because the efficiencies have to be positive.

Table 6 contains the estimation results from the efficiency correction model for 67 local exchange carriers from 1997 until 2006. A satisfying aspect is that the measurement errors are low: $\sigma_\varepsilon = 0.021$, and the degrees of freedom ν are approximately 6. Inspection of the data shows that rather violent changes took place during the sample period. Mergers and takeovers took place after the liberalization of the market for LEC 's in 1996 and huge technological shifts occurred.

The estimates of β reveal that only a simple model remains: economies of scale hardly above 1, a clear influence of SH^*, and some influence of LL^*. The specification error has a standard error $\sigma_v = 0.16$. Given the huge technological shifts and differences in circumstances in the various states of the USA this standard error seems reasonable.

The estimate of γ_t show a decrease in cost level during the first two years followed by a steady increase. The linear trend κ is 0.036 with a relatively large standard error of 0.017. The standard error $\sigma_\zeta = 0.048$ is also high. Obviously much has been going on in the sample period, the yearly changes are sizable and irregular. If the random walk with drift would be replaced by fixed dummy variables, the results for γ_t would hardly change.

The estimates of δ_1 and δ_2 are clearly different from 1 and have low standard errors. It is difficult to explain intuitively that δ_2 is so much larger than δ_1.

The covariance stationarity restrictions make that the results for (ϕ, λ) follow directly from those for (ϕ_1, λ_1). The relevant aspects of the estimates are the moments for u_{i2}. The expectation of u_{i2}, ϕ_1/λ_1, is very well identified: 0.155 with a standard error of only 0.009. The standard error, $\sqrt{\phi_1}/\lambda_1$ too: 0.072 with a standard error of 0.006. Imposing the restriction was sufficient to obtain a well identified model.

The DIC criterion cannot be compared directly to that of the error correction random effects model (19)–(21) in table 4, because the latter results are based on the whole sample. Estimation of ECREM on the dataset without the first year gives a DIC of -1911.2. The efficiency correction model leads to a gain in DIC of 617 points. The estimated effective number of parameters p_D however is relatively high, 490, a large proportion of the 670 observations.

Table 7 gives the posterior means of the inefficiency levels u_{it} and the specification errors v_i for all local exchange carriers. As explained the nominal values of the u_{it} have little statistical meaning. The autocorrelations of the u_{it} are as expected: the correlation between u_{i2} and u_{it} is around $(1-\delta_1)^{t-2}$. Further the means of the inefficiencies over time $\bar{u}_{i.}$ and the specification errors v_i are virtually uncorrelated.

To check the sensitivity of the assumption $P(u_{i2} < 0.05) = 0.02$ on the estimated inefficiencies we replaced the assumption by $P(u_{i2} < 0.01) = 0.02$, approximated by $\lambda_1 = -20.1 + 21\phi_1$. The main change in the results is a shift in $\mathrm{E}(u_{i12})$ from 0.155 to 0.091. The robustness of the outcomes was confirmed by the correlation between the the posterior means of the estimated inefficiencies u_{it} in the two approaches, being 0.99. The estimated relative efficiencies are virtually equal. It may be concluded that model provides robust estimates of the relative inefficiencies.

6 APPLICATION TO HOSPITAL DATA

In this section the same analysis as done for the LEC's is applied to the dataset of 382 hospitals over a time period of 5 years as analyzed in Koop, Osiewalski, and Steel (1997), Griffin and Steel (2004, 2007) and Atkinson and Dorfman (2005). These studies do not use firm specific stochastic dynamic efficiency changes. The specification in Griffin and Steel (2004) assumes time invariant inefficiency u_i, unobserved heterogeneity is not included, measurement errors are. A vague prior is used for the distribution of u_i, a Bayesian nonparametric method is used for estimation of u_i. Atkinson and Dorfman (2005) use a model with a deterministic specification for firm specific time varying inefficiency. For estimation they apply the Bayesian method of moments in a Gibbs sampling framework.

In this section a simplified model without interaction effects is analyzed. The variables are provided in table 3. The variable D is a scaling variable, so it's coefficient is the economy of scale parameter.

Table 8 provides the estimation results from the preliminary analysis as in section 4. The results are like those for the LEC's: the economy of scale parameter D is 0.99 for the model in means (17), and 0.74 for the model in deviations (18). The results from REM (16) are thus based on a hypothesis that should be refuted.

Table 9 contains the estimation results from the error correction random effects models (19)–(21). They clearly perform better than the standard random effects model in table 8. The loglikelihood gain by the introduction of δ_1 and δ_2 is 235 points. In the model without restrictions (the third panel of table 8) both δ_1 and δ_2 significantly differ from 1: the estimates of δ_1 and δ_2 are respectively 0.30 and 0.59, both with small standard errors.

The Bayesian estimation results from the error correction random effects models by WinBUGS are almost a copy of the maximum likelihood results. For the case $\delta_2 = 1$ the DIC outcome is -4381 and the effective number of parameters p_D equals 267. This result can be compared to the "best" model in Griffin and Steel (2007), who also use WinBUGS for inference on the efficiency of hospitals. In a model with deterministic time varying inefficiency and no specification error they obtain a DIC of 4834 with an effective number of parameters of 398. They use 28 additional explanatory variables (cross products) which clearly improves the fit. For convenience we stick to our simple model. When the observations from the first year are excluded in ECREM with $\delta_2 = 1$, the DIC and effective number of parameters p_D equal -3200 and 141, respectively. These results can be used to compare the outcomes from the efficiency correction model.

Note that the high values of p_D are due to the random effects. The penalty is a fraction of the number of firms N, where the fraction depends on σ_v. In case σ_v is large, the random effects model approaches a fixed effects model, resulting in a penalty N.

Table 10 provides the estimation results from the efficiency correction model (11)–(13). As the number of time periods is only 4, the time dependence is simply modeled by dummy variables instead of a random walk with drift. The estimates of the time dummy variables reflect a steady cost increase.

The crucial parameters δ_1 (0.358) and δ_2 (0.617) have very low standard errors and are clearly different from 1 and from each other, and they are quite close to the estimates from the error correction random effects model. The DIC gain is more than 2500 points, however the p_D is extremely large, 1376, corresponding to 95% of the number of observations.

The economies of scale parameter D is as expected approximately 1. All regression coefficients are very significant. The estimate of σ_ε is only 0.014, but the very low value of the degrees of freedom for the student distribution, $\nu = 2.2$, suggests that there are some severe outliers. The estimate of σ_v (0.11) is very acceptable, and could even be lower when more explanatory variables were used.

The inefficiencies have an expectation of 0.135 and a standard error of 0.059. Both values are slightly lower than in the case of the LEC's. The similarity is of course mainly the result of the imposed restriction.

7 CONCLUSIONS

The efficiency correction model performs well for local exchange carriers as well as hospitals. It is a clear improvement over existing models in terms of plausibility as well as statistical fit. The similarity of the outcomes for two such different sectors suggests general applicability.

The crucial question is whether this model provides sufficient information to judge the inefficiency of companies, like for example NERA intended to do. We got robust estimates for relative inefficiencies within each year under the assumption that specification errors are time invariant. This assumption might be questioned, but that raises difficult identification issues. Moreover we assumed that there are no time invariant inefficiencies. If this assumption is incorrect, time invariant inefficiencies are part of the specification error v_i. In theory this could be investigated by decomposing v_i in a symmetric specification error and a positive skewed time invariant inefficiency term per firm. With noninformative priors it is even theoretically problematic to make this distinction, as shown by Fernández, Osiewalski, and Steel (1997). Strong informative priors will give a result, but the posteriors will hardly differ from the priors.

It will be hard to argue what the situation is for a firm with a negative value for v_i (an unexplained time invariant low cost level) combined with a high inefficiency u_{it}. Developing better models to reduce the specification error σ_v seems the only way out.

References

Ahn, S. C., D. H. Good, and R. C. Sickles. 2000. "Estimation of Long-Run Inefficiency Levels: A Dynamic Frontier Approach." *Econometric Theory* 19:461–492.

Atkinson, S. E., and J. H. Dorfman. 2005. "Bayesian Measurement of Productivity and Efficiency in the Presence of Undesirable Outputs: Crediting Electric Utilities for Reducing Air Pollution." *Journal of Econometrics* 126:445–468.

Battesse, 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* 3:153–169.

Cornwell, C., P. Schmidt, and R. C. Sickles. 1990. "Production Frontiers with Cross-Sectional and Time-Series Variation in Efficiency Levels." *Journal of Econometrics* 46:185–200.

Desli, E., S. C. Ray, and S. C. Kumbhakar. 2003. "A Dynamic Stochastic Frontier Production Model with Time-Varying Efficiency." *Applied Economics Letters* 10:623–626.

Farsi, M., M. Filippini, and W. Greene. 2006. "Application of Panel Data Models in Benchmarking Analysis of the Electricity Distribution Sector." *Annals of Public and Cooperative Economics* 77:271–290.

Fernández, C., J. Osiewalski, and M. F. J. Steel. 1997. "On the Use of Panel Data in Stochastic Frontier Models with Improper Priors." *Journal of Econometrics* 79:169–193.

Greene, W. 2005. "Reconsidering heterogeneity in panel data estimators of the stochastic frontier model." *Journal of Econometrics* 126:269–303.

Griffin, J. E., and M. F. J. Steel. 2004. "Semiparametric Bayesian Inference for stochastic Frontier Models." *Journal of Econometrics* 123:121–152.

———. 2007. "Bayesian Stochastic Frontier analysis using WinBUGS." *Journal of Productivity Analysis* 27:163–176.

Grunwald, G. K., R. J. Hyndman, L. Tedesco, and R. L. Tweedie. 2000. "Non-Gaussian Conditional Linear AR(1) Models." *Australian & New Zealand Journal of Statistics* 42:479–495.

Koop, G., J. Osiewalski, and M.F.J. Steel. 1997. "Bayesian Efficiency Analysis through individual effects: Hospital cost frontiers." *Journal of Econometrics* 76:77–105.

Kumbhakar, S.C., and E.G. Tsionas. 2005. "Measuring Technical and Allocative Inefficiency in the Translog Cost System: a Bayesian Approach." *Journal of Econometrics* 126:355–384.

NERA Economic Consulting. 2005. "The Comparative Efficiency of BT in 2003: A Report for Ofcom." Technical Report.

———. 2006. "The Comparative Efficiency of KPN; a Report for OPTA." Technical Report.

Park, B. U., R. C. Sickles, and L. Simar. 2003. "Semiparametric-Efficient Estimation of AR(1) Panel Data Models." *Journal of Econometrics* 117:279–303.

———. 2007. "Semiparametric Efficient Estimation of Dynamic Panel Data Models." *Journal of Econometrics* 136:281–301.

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. van der Linde. 2002. "Bayesian measures of model complexity and fit." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64 (4): 583–639.

Tsionas, E. G. 2002. "Stochastic Frontier Models with Random Coefficients." *Journal of Applied Econometrics* 17:127–147.

———. 2006. "Inference in Dynamic Stochastic Frontier Models." *Journal of Applied Econometrics* 21:669–676.

Tsionas, E. G., and S. C Kumbhakar. 2004. "Markov Switching Stochastic Frontier Model." *Econometrics Journal* 7:398–425.

Variable	Description
C	Costs
SL	Switched lines
LL	Leased Lines
SM	Switch Minutes
SH	Sheath
PD	Population Density
BR	Business-to-residential ratio
D*t*	Dummy variable for year *t*

Table 1: Definition of LEC variables.

Year	C (x 1,000)	SL (x 1,000)	LL (x 1,000)	SM (x 1,000)	SH (x 1,000)	RB	PD
1996	884.7	2,106.2	523.1	40,327.5	77.9	0.41	65.8
1997	900.9	2,213.6	663.0	42,900.5	79.1	0.43	66.4
1998	911.2	2,303.4	917.2	47,072.7	80.6	0.45	67.1
1999	955.2	2,379.2	1,375.1	47,806.6	81.8	0.46	67.7
2000	986.6	2,377.8	1,770.7	47,606.1	81.9	0.47	68.4
2001	1,007.2	2,265.1	2,122.3	43,875.1	83.1	0.45	69.0
2002	1,035.5	2,149.1	2,463.9	36,635.8	84.5	0.48	69.6
2003	1,076.1	2,006.0	2,758.4	31,863.7	85.3	0.47	70.2
2004	1,052.6	1,904.3	3,029.5	30,008.3	86.6	0.48	70.7
2005	1,116.7	1,824.5	4,419.2	26,412.4	89.1	0.49	71.0
2006	1,138.1	1,723.5	5,128.7	23,160.0	90.4	0.52	71.4

Table 2: Averages of the cost, output and environmental variables for the LEC's.

Variable	Description
D	Number of inpatient days
C	Number of cases
B	Number of beds
O	Number of outpatient visits
CMI	Case mix Index
AWI	Aggregate wage index
CS	Capital stock
D*t*	Dummy variable for year *t*

Table 3: Definition of hospital variables.

	Means (OLS)			Deviations (OLS)			REM (ML)			REM (Bayes)			
	Coef	Sd	T-val	Coef	Sd	T-val	Coef	Sd	T-val	Mean	sd	2.5%	97.5%
Const	-0.248	0.307	-0.81	4.451	0.549	8.11	0.941	0.283	3.33	1.125	0.443	0.353	2.087
ln(SL)	0.982	0.019	51.78	0.592	0.048	12.27	0.963	0.025	38.31	0.948	0.039	0.863	1.017
ln(LL)*	0.059	0.053	1.11	0.048	0.007	6.87	0.043	0.007	6.00	0.044	0.007	0.030	0.059
ln(SM)*	0.041	0.060	0.68	0.020	0.011	1.82	0.011	0.011	1.02	0.008	0.012	-0.015	0.030
ln(SH)*	0.055	0.051	1.07	0.418	0.043	9.68	0.533	0.036	14.69	0.533	0.037	0.458	0.606
ln(PD)	-0.014	0.017	-0.84	0.441	0.089	4.98	0.105	0.029	3.64	0.115	0.036	0.050	0.193
ln(BR)	0.162	0.115	1.41	-0.084	0.028	-2.98	-0.009	0.028	-0.34	-0.011	0.030	-0.069	0.048
D97				-0.037	0.010	-3.59	-0.048	0.011	-4.50	-0.047	0.011	-0.069	-0.026
D98				-0.050	0.011	-4.40	-0.071	0.011	-6.23	-0.071	0.012	-0.093	-0.048
D99				-0.068	0.013	-5.24	-0.097	0.013	-7.60	-0.097	0.014	-0.123	-0.070
D00				-0.057	0.014	-4.03	-0.083	0.014	-6.03	-0.082	0.014	-0.110	-0.055
D01				-0.063	0.015	-4.24	-0.071	0.014	-4.93	-0.072	0.014	-0.099	-0.044
D02				-0.042	0.016	-2.54	-0.042	0.016	-2.65	-0.043	0.016	-0.075	-0.013
D03				0.003	0.018	0.15	0.024	0.017	1.44	0.021	0.017	-0.012	0.055
D04				-0.023	0.020	-1.15	0.014	0.018	0.73	0.010	0.019	-0.028	0.047
D05				0.032	0.023	1.39	0.081	0.021	3.85	0.073	0.022	0.029	0.115
D06				0.032	0.025	1.26	0.095	0.023	4.07	0.088	0.024	0.041	0.134
$\sqrt{\frac{\sigma_\alpha^2}{T}+\sigma_\theta^2}$	0.121												
σ_α				0.057			0.060			0.060	0.002	0.056	0.063
σ_θ							0.271			0.294	0.044	0.221	0.392
LL							848.9						
DIC										-1990.5			
p_D										82.5			

Table 4: Estimation results from the model in means (17), deviations (18), and the random effects model (REM) (16) for the LEC data.

	$\delta_2 = 1$			$\delta_1 = \delta_2$			No restriction on δ			$\delta_2 = 1$ (Bayes)			
	Coef	Sd	T-val	Coef	Sd	T-val	Coef	Sd	T-val	Mean	Sd	2.5%	97.5%
Const	-0.287	0.176	-1.63	-0.262	0.245	-1.07	-0.409	0.206	-1.99	-0.235	0.198	-0.615	0.161
ln(SL)	1.003	0.015	66.70	0.982	0.016	60.54	0.989	0.016	63.72	1.004	0.017	0.970	1.037
ln(LL)*	0.031	0.010	3.14	0.053	0.026	2.08	0.039	0.019	2.03	0.030	0.010	0.011	0.049
ln(SM)*	0.019	0.014	1.34	0.010	0.038	0.27	0.027	0.028	0.98	0.018	0.014	-0.009	0.046
ln(SH)*	0.203	0.037	5.55	0.075	0.046	1.62	0.087	0.042	2.05	0.233	0.044	0.152	0.322
ln(PD)	0.006	0.016	0.38	-0.001	0.017	-0.08	-0.004	0.016	-0.27	0.012	0.018	-0.022	0.047
ln(BR)	-0.017	0.033	-0.50	-0.001	0.087	-0.01	-0.026	0.063	-0.42	-0.021	0.033	-0.085	0.043
D97	-0.058	0.007	-7.96	-0.122	0.046	-2.65	-0.075	0.016	-4.78	-0.056	0.007	-0.070	-0.042
D98	-0.087	0.011	-7.83	-0.002	0.048	-0.04	-0.083	0.022	-3.69	-0.084	0.011	-0.105	-0.063
D99	-0.118	0.015	-8.00	-0.068	0.051	-1.33	-0.104	0.029	-3.62	-0.113	0.014	-0.141	-0.085
D00	-0.098	0.017	-5.72	0.011	0.054	0.20	-0.058	0.033	-1.78	-0.094	0.017	-0.126	-0.061
D01	-0.065	0.019	-3.38	-0.035	0.057	-0.62	-0.019	0.037	-0.51	-0.062	0.018	-0.099	-0.025
D02	-0.008	0.022	-0.37	0.137	0.061	2.26	0.074	0.040	1.83	-0.007	0.021	-0.047	0.034
D03	0.088	0.024	3.75	0.357	0.064	5.57	0.226	0.044	5.19	0.087	0.022	0.045	0.132
D04	0.104	0.026	4.04	-0.038	0.068	-0.55	0.196	0.047	4.18	0.101	0.025	0.054	0.151
D05	0.202	0.029	6.94	0.593	0.076	7.86	0.377	0.053	7.10	0.196	0.028	0.140	0.252
D06	0.247	0.032	7.70	0.244	0.082	2.98	0.403	0.058	6.94	0.241	0.032	0.180	0.303
δ_1	0.080	0.015	5.43	0.143	0.035	4.12	0.078	0.012	6.53	0.086	0.024	0.048	0.141
δ_2	1			0.143	0.035	4.12	0.429	0.079	5.46	1		0.048	0.053
σ_η	0.053			0.051			0.051			0.051	0.001	0.057	0.156
σ_θ	0.072			0.096			0.052			0.103	0.025		
LL	1046.8			1067.3			1077.1						
DIC										-2142.4			
p_D										49.3			

Table 5: Estimation results from the error correction random effects model (ECREM) (19)–(21) for the LEC data.

Variable	Mean	Sd	2.5%	97.5%
Const	-0.674	0.233	-1.116	-0.186
ln(SL)	1.036	0.019	0.998	1.072
ln(LL)*	0.029	0.010	0.010	0.047
ln(SM)*	0.003	0.018	-0.033	0.037
ln(SH)*	0.294	0.055	0.193	0.407
ln(PD)	0.015	0.021	-0.025	0.060
ln(BR)	-0.021	0.038	-0.094	0.055
D98	-0.025	0.008	-0.040	-0.010
D99	-0.049	0.011	-0.071	-0.028
D00	-0.031	0.014	-0.059	-0.004
D01	0.003	0.016	-0.028	0.034
D02	0.054	0.019	0.018	0.091
D03	0.136	0.022	0.093	0.179
D04	0.155	0.024	0.107	0.203
D05	0.239	0.028	0.184	0.294
D06	0.289	0.032	0.226	0.352
δ_1	0.261	0.031	0.206	0.326
δ_2	0.756	0.069	0.619	0.890
σ_ε	0.021	0.002	0.017	0.026
ν	6.058	2.468	3.159	12.63
κ	0.036	0.017	0.001	0.069
σ_ζ	0.048	0.014	0.029	0.082
σ_v	0.163	0.021	0.127	0.212
ϕ_1	4.653	0.228	4.211	5.105
ϕ	0.702	0.117	0.503	0.956
λ_1	30.26	31.88	24.32	36.80
λ	17.42	20.09	13.69	21.54
ϕ_1/λ_1	0.155	0.009	0.139	0.173
ϕ/λ	0.040	0.004	0.033	0.049
$\sqrt{\phi_1}/\lambda_1$	0.072	0.006	0.061	0.084
$\sqrt{\phi}/\lambda$	0.048	0.003	0.042	0.055
p_D	490.1			
DIC	-2528.3			

Table 6: Estimation results from the efficiency correction model (11)–(13) and (22) for the LEC data, where $P(u_{i1} < .05) = .02$.

LEC	$u_{i,97}$	$u_{i,98}$	$u_{i,99}$	$u_{i,00}$	$u_{i,01}$	$u_{i,02}$	$u_{i,03}$	$u_{i,04}$	$u_{i,05}$	$u_{i,06}$	$\bar{u}_{i,.}$	v_i
1	0.1393	0.1414	0.1572	0.1564	0.1316	0.1336	0.1243	0.1009	0.0886	0.0869	0.1260	0.0381
2	0.1374	0.2034	0.1869	0.1539	0.1603	0.1728	0.1435	0.1350	0.1889	0.1514	0.1634	0.1938
3	0.1520	0.2113	0.1863	0.1483	0.1467	0.1864	0.1635	0.1536	0.1343	0.1095	0.1592	0.0226
4	0.1201	0.1293	0.1291	0.1183	0.1323	0.1155	0.0991	0.0862	0.0934	0.0975	0.1121	0.1158
5	0.1698	0.1899	0.1643	0.1399	0.1159	0.1135	0.1398	0.1283	0.1099	0.0917	0.1363	0.0244
6	0.0931	0.1029	0.1686	0.1738	0.1460	0.1179	0.1097	0.1348	0.1640	0.1873	0.1398	-0.0002
7	0.1917	0.1675	0.1309	0.1110	0.0940	0.0852	0.1844	0.1717	0.1617	0.1459	0.1444	-0.3553
8	0.7123	0.5416	0.4161	0.3200	0.2513	0.3308	0.2940	0.2737	0.4498	0.3464	0.3936	-0.0198
9	0.2962	0.2452	0.1890	0.1500	0.1405	0.1219	0.1342	0.1262	0.1588	0.1368	0.1699	0.0065
10	0.0903	0.1298	0.1241	0.1227	0.1045	0.1067	0.1356	0.1199	0.1446	0.1841	0.1262	0.1137
11	0.2919	0.2778	0.2184	0.1720	0.1588	0.1691	0.1788	0.1568	0.1330	0.1082	0.1865	0.0341
12	0.1954	0.1711	0.1333	0.1065	0.0952	0.1029	0.1604	0.1615	0.1663	0.1413	0.1434	-0.3809
13	0.2493	0.1979	0.1566	0.1240	0.1096	0.1041	0.1297	0.1347	0.1562	0.1356	0.1498	-0.1532
14	0.3121	0.3339	0.2744	0.2188	0.2212	0.1971	0.1817	0.1440	0.1138	0.0937	0.2091	-0.3693
15	0.2587	0.2229	0.2046	0.2021	0.1671	0.1474	0.2271	0.1794	0.1470	0.1187	0.1875	0.1092
16	0.1811	0.1921	0.1516	0.1233	0.1088	0.1009	0.1052	0.0962	0.0862	0.0785	0.1224	-0.3256
17	0.1079	0.1023	0.0894	0.0822	0.1031	0.1659	0.1691	0.1404	0.1709	0.1788	0.1310	0.0930
18	0.1622	0.2032	0.1684	0.1379	0.1175	0.1149	0.1696	0.1427	0.1173	0.0976	0.1431	-0.2680
19	0.2005	0.1610	0.1278	0.1172	0.0980	0.1116	0.1701	0.1949	0.1762	0.2253	0.1583	0.0168
20	0.0833	0.0771	0.0749	0.1042	0.1247	0.1286	0.1672	0.1630	0.1975	0.1943	0.1315	0.1006
21	0.3438	0.3083	0.3384	0.3264	0.2563	0.2009	0.1643	0.1299	0.1088	0.0986	0.2276	-0.2282
22	0.2094	0.1711	0.1396	0.1218	0.1081	0.1294	0.1276	0.1128	0.1023	0.0907	0.1313	0.1464
23	0.1364	0.1763	0.1429	0.1248	0.1454	0.1282	0.1442	0.1298	0.1075	0.0905	0.1326	-0.2638
24	0.0900	0.0801	0.1024	0.1418	0.1640	0.2374	0.2431	0.2082	0.1674	0.1532	0.1588	0.0127
25	0.0743	0.0697	0.0718	0.0819	0.1039	0.2242	0.2478	0.2601	0.2230	0.2329	0.1590	-0.0882
26	0.0929	0.1021	0.1724	0.2167	0.1790	0.1473	0.1611	0.1991	0.1743	0.1535	0.1598	0.1397
27	0.1426	0.1547	0.2643	0.2944	0.2328	0.1850	0.1661	0.1471	0.1394	0.1171	0.1844	0.2188
28	0.2349	0.2073	0.1883	0.1623	0.1342	0.1358	0.1137	0.0920	0.0804	0.0757	0.1425	-0.2078
29	0.3659	0.2999	0.2386	0.1948	0.1544	0.1220	0.1050	0.1223	0.1111	0.1104	0.1824	-0.2259
30	0.2297	0.2463	0.2205	0.1912	0.1518	0.1212	0.1032	0.1033	0.0985	0.0933	0.1559	0.0133
31	0.0843	0.0926	0.1362	0.1517	0.1530	0.1646	0.1433	0.1394	0.1393	0.1216	0.1326	0.2043
32	0.2759	0.2561	0.2148	0.1881	0.1556	0.1280	0.1145	0.0991	0.0834	0.0824	0.1598	0.0760
33	0.0855	0.0741	0.0864	0.1060	0.0930	0.1391	0.1821	0.1961	0.1872	0.1772	0.1327	-0.1129
34	0.1177	0.1245	0.1095	0.1011	0.0995	0.1366	0.1937	0.1639	0.1377	0.1528	0.1337	0.2841
35	0.2001	0.3108	0.2439	0.1951	0.1571	0.1274	0.1184	0.1138	0.1127	0.1088	0.1688	-0.0377
36	0.1067	0.1639	0.1385	0.1182	0.1415	0.1428	0.1314	0.1110	0.1579	0.1603	0.1372	0.1127
37	0.0844	0.1250	0.1191	0.1281	0.1346	0.1124	0.1796	0.1936	0.2467	0.2780	0.1601	0.0591
38	0.3419	0.3673	0.3445	0.3349	0.3059	0.2367	0.1821	0.1448	0.1358	0.1597	0.2554	-0.0767
39	0.1547	0.1751	0.2027	0.1870	0.1614	0.1528	0.1527	0.1212	0.1004	0.0929	0.1501	0.0565
40	0.1347	0.1218	0.1166	0.1160	0.1073	0.0909	0.0787	0.0855	0.0775	0.0870	0.1016	-0.1478
41	0.0745	0.0725	0.0953	0.1760	0.1545	0.1366	0.1596	0.1783	0.1724	0.1743	0.1394	-0.0280
42	0.1930	0.1544	0.1226	0.1090	0.1289	0.1078	0.1435	0.2029	0.2445	0.2589	0.1666	-0.0681
43	0.1655	0.1632	0.1770	0.1433	0.1167	0.0986	0.2215	0.2675	0.3547	0.3125	0.2021	0.2197
44	0.0872	0.0791	0.0903	0.1417	0.1239	0.1134	0.1088	0.2020	0.2456	0.2313	0.1423	0.0250
45	0.1082	0.1323	0.1987	0.1739	0.1758	0.1735	0.1975	0.1715	0.1391	0.1383	0.1609	0.3914
46	0.0897	0.0763	0.0791	0.1249	0.1148	0.1722	0.2401	0.2141	0.1836	0.1797	0.1474	-0.0687
47	0.1306	0.1475	0.1745	0.1783	0.1477	0.1208	0.1398	0.1191	0.1022	0.1040	0.1365	0.0385
48	0.1388	0.1247	0.1333	0.1202	0.1010	0.0926	0.1050	0.1275	0.1153	0.1014	0.1160	0.1563
49	0.1887	0.1832	0.2143	0.2053	0.1664	0.1323	0.1221	0.1154	0.0982	0.0911	0.1517	0.1031
50	0.1443	0.1205	0.1794	0.1660	0.1808	0.1663	0.1339	0.1154	0.1082	0.0976	0.1412	0.0489
51	0.3358	0.3751	0.4033	0.4093	0.3798	0.3451	0.2658	0.2048	0.1615	0.1375	0.3018	-0.0655
52	0.1439	0.1168	0.1041	0.0964	0.1449	0.1357	0.1130	0.1562	0.1615	0.1461	0.1319	0.0386
53	0.0996	0.0833	0.0843	0.1328	0.1980	0.2682	0.2300	0.2641	0.2790	0.2354	0.1875	0.1119
54	0.1289	0.1090	0.1047	0.1201	0.1376	0.1153	0.0962	0.1009	0.1294	0.1271	0.1169	0.0511
55	0.1425	0.1156	0.1006	0.0971	0.0924	0.1011	0.0874	0.1125	0.1347	0.1349	0.1119	0.0010
56	0.1847	0.1523	0.1268	0.1095	0.0944	0.0938	0.0832	0.1292	0.1762	0.1572	0.1307	0.0297
57	0.1393	0.1140	0.1018	0.0977	0.0846	0.0871	0.0874	0.1433	0.1412	0.1281	0.1124	0.0009
58	0.1422	0.1150	0.1047	0.0969	0.0934	0.0941	0.0833	0.1344	0.2821	0.2475	0.1394	-0.0022
59	0.1517	0.1212	0.1027	0.1084	0.1065	0.1165	0.0989	0.1458	0.2814	0.2393	0.1472	-0.0457
60	0.1280	0.1066	0.1055	0.0915	0.1111	0.1040	0.0919	0.1057	0.1075	0.1027	0.1055	-0.0492
61	0.1992	0.2105	0.2093	0.2409	0.2059	0.1772	0.1387	0.1146	0.1018	0.1259	0.1724	0.1216
62	0.1506	0.1598	0.2299	0.2912	0.2604	0.2062	0.1849	0.1641	0.1296	0.1096	0.1886	-0.0168
63	0.1657	0.1418	0.1544	0.1232	0.1088	0.1714	0.1534	0.1314	0.1188	0.1095	0.1378	0.0309
64	0.1823	0.1706	0.1642	0.1383	0.1463	0.1547	0.1253	0.1041	0.0940	0.0886	0.1368	0.0987
65	0.1896	0.2159	0.2653	0.2187	0.1877	0.1779	0.1411	0.1128	0.1044	0.0966	0.1710	-0.0776
66	0.1142	0.1048	0.2007	0.2015	0.2450	0.2317	0.1868	0.1630	0.1359	0.1180	0.1702	0.1625
67	0.0698	0.0650	0.1363	0.1295	0.1585	0.2382	0.2619	0.2579	0.2714	0.2922	0.1881	-0.1411
$\bar{u}_{.,t}$	0.1737	0.1695	0.1673	0.1598	0.1497	0.1496	0.1528	0.1504	0.1556	0.1467		
γ_t	0	-0.0246	-0.0492	-0.0309	0.0026	0.0535	0.1364	0.1552	0.2393	0.2891		

Table 7: Efficiencies of the LEC's per year in the efficiency correction model (EFCOM) (11)–(13), where $P(u_{i1} < .05) = .02$.

	Means (OLS)			Deviations (OLS)			REM (ML)		
	Coef	Sd	T-val	Coef	Sd	T-val	Coef	Sd	T-val
Const	7.242	0.191	37.83	8.773	0.268	32.77	6.705	0.131	51.31
D	0.991	0.017	58.82	0.740	0.029	25.27	0.981	0.013	76.13
C*	0.292	0.033	8.85	0.186	0.018	10.37	0.232	0.016	14.34
B*	0.166	0.040	4.16	0.001	0.024	0.03	0.114	0.020	5.68
O*	0.037	0.014	2.71	0.037	0.007	5.48	0.038	0.006	6.18
CMI	0.943	0.091	10.34	0.032	0.050	0.63	0.353	0.044	7.94
AWI	0.701	0.039	17.86	0.221	0.064	3.46	0.614	0.035	17.78
CS*	0.100	0.013	7.43	0.103	0.010	10.60	0.126	0.008	15.61
D88				0.100	0.006	16.69	0.113	0.005	20.91
D89				0.199	0.007	30.47	0.206	0.006	35.01
D90				0.296	0.007	43.39	0.298	0.006	48.30
D91				0.390	0.007	52.36	0.386	0.007	57.17
$\sqrt{\frac{\sigma_\alpha^2}{T}+\sigma_\theta^2}$	0.118								
σ_α				0.062			0.065		
σ_θ							0.115		
LL							1975.9		

Table 8: Estimation results from the model in means (17), deviations (18), and the random effects model (REM) (16) for the hospital data.

	$\delta_2 = 1$			$\delta_1 = \delta_2$			No restriction on δ		
	Coef	Sd	T-val	Coef	Sd	T-val	Coef	Sd	T-val
Const	6.715	0.136	49.45	6.240	0.154	40.40	6.198	0.154	40.20
D	0.977	0.013	74.76	1.002	0.014	70.81	1,000	0.014	71.02
C*	0.199	0.016	12.79	0.287	0.023	12.37	0.267	0.022	12.29
B*	0.126	0.020	6.25	0.056	0.027	2.06	0.057	0.026	2.15
O*	0.027	0.006	4.34	0.029	0.009	3.36	0.024	0.008	2.84
CMI	0.305	0.046	6.61	0.457	0.063	7.22	0.412	0.062	6.65
AWI	0.578	0.035	16.61	0.665	0.038	17.67	0.641	0.037	17.20
CS*	0.133	0.008	15.83	0.128	0.010	12.28	0.135	0.010	13.11
D88	0.113	0.005	24.33	0.212	0.008	25.69	0.179	0.007	26.03
D89	0.208	0.006	34.04	0.315	0.009	35.32	0.311	0.008	36.63
D90	0.301	0.007	43.76	0.414	0.009	44.67	0.427	0.009	46.46
D91	0.390	0.008	50.47	0.479	0.010	47.76	0.512	0.010	50.86
δ_1	1			0.490	0.023	21.08	0.301	0.035	8.56
δ_2	0.329	0.042	7.89	0.490	0.023	21.08	0.592	0.028	21.30
σ_η	0.067			0.063			0.064		
σ_θ	0.098			0.105			0.093		
LL	2118.3			2185.6			2210.2		

Table 9: Estimation results from the error correction model (ECREM) (19)–(21) for the hospital data.

Variable	Mean	Sd	2.5%	97.5%
Const	6.576	0.148	6.286	6.866
D	0.987	0.014	0.960	1.015
C*	0.205	0.022	0.163	0.246
B*	0.118	0.024	0.071	0.164
O*	0.041	0.007	0.027	0.055
CMI	0.494	0.054	0.389	0.600
AWI	0.579	0.036	0.509	0.649
CS*	0.121	0.010	0.101	0.141
D89	0.107	0.003	0.101	0.114
D90	0.200	0.004	0.192	0.209
D91	0.284	0.005	0.274	0.295
δ_1	0.358	0.033	0.295	0.422
δ_2	0.617	0.024	0.571	0.663
σ_ε	0.014	0.002	0.011	0.017
ν	2.162	0.242	1.745	2.688
σ_v	0.111	0.005	0.101	0.121
ϕ_1	5.222	0.188	4.863	5.592
ϕ	1.144	0.153	0.864	1.462
λ_1	38.68	2.966	33.18	44.67
λ	23.59	2.098	19.77	27.83
ϕ_1/λ_1	0.135	0.005	0.125	0.147
ϕ/λ	0.048	0.004	0.041	0.056
$\sqrt{\phi_1}/\lambda_1$	0.059	0.003	0.053	0.066
$\sqrt{\phi}/\lambda$	0.045	0.002	0.041	0.050
p_D	1375.8			
DIC	-5794.0			

Table 10: Estimation results from the efficiency correction model (11)–(13) for the hospital data, where $P(u_{i1} < .05) = .02$.