Measuring Productivity and Efficiency: A Kalman Filter Approach

Meryem Duygun

Nottingham University Business School, Nottingham University

Levent Kutlu

School of Economics, Georgia Institute of Technology

Department of Economics, Antalya International University

Robin C. Sickles

Department of Economics, Rice University

July 31, 2016

Abstract

In the Kalman filter setting, one can model the inefficiency term of the standard stochastic frontier composed error as an unobserved state. In this study a panel data version of the local level model is used for estimating time-varying efficiencies of firms. We apply the Kalman filter to estimate average efficiencies of U.S. airlines and find that the technical efficiency of these carriers did not improve during the period 1999-2009. During this period the industry incurred substantial losses, and the efficiency gains from reorganized networks, code-sharing arrangements, and other best business practices apparently had already been realized.

Keywords: Kalman filter, panel data, airline productivity

JEL classification numbers: C13, C23, R49

1 Introduction

Stochastic frontier analysis originated with two seminal papers, Meeusen and van den Broeck (1977) and Aigner, Lovell, and Schmidt (1977). Jondrow et al. (1982) provided a way to estimate firm specific technical efficiency. These contributions were framed in a cross sectional data framework. Panel data potentially can give more reliable information about the efficiencies of the firm. Pitt and Lee (1981) and Schmidt and Sickles (1984) applied random effects and fixed effects models to estimate firm specific efficiencies. In these models the efficiencies are assumed to be time-invariant. For long panel data this assumption might be questionable. The time-invariance assumption was relaxed by Cornwell, Schmidt, and Sickles (1990) (CSS), Kumbhakar (1990), Battese and Coelli (1992) (BC), and Lee and Schmidt (1992). The time-varying inefficiency models were followed by dynamic efficiency models such as Ahn, Good, and Sickles (2000), Desli, Ray, and Kumbhakar (2003), Tsionas (2006), Huang and Chen (2009), and Assaf, Gillen, and Tsionas (2014).[1] Work on time varying effects models and their use in productivity and efficiency studies have accelerated in the last decade and we view our current contribution as following in this tradition. Many of these advances are summarized in the recent chapter by Sickles, Hao, and Sheng (2015).

In this paper we consider the use of the Kalman (1960) filter by treating the inefficiency term as an unobserved state. In contrast to the classical Box-Jenkins approach, one also can explicitly model non-stationary stochastic processes in the Kalman filter setting. This gives significant flexibility to the econometrician when specifying the inefficiency portion of the model. We use the Kalman filter estimator (KFE) to model the efficiency component of the stochastic frontier composed error. For this purpose we use a panel data generalization of the local level model. For long panel data, relatively inflexible

---

[1]See also Galán and Pollitt (2014) for an empirical study.

stochastic frontier models (e.g., BC, CSS, and Kumbhakar (1990)) are more likely to fail to capture potentially complex time-varying patterns of the effects terms. We examine this claim by conducting a series of Monte Carlo simulations. Results of these simulations indicate that some of the widely used estimators can perform poorly in terms of capturing the efficiencies of firms when we have long panel data with fluctuating efficiencies. For example, if the efficiencies of firms are affected by macro factors that tend to have cycles, then it is likely that these relatively inflexible approximations will fail to capture the efficiency patterns. While some of the factors that lead to variation in efficiency can be controlled for by including exogenous variables in the modeling of the inefficiency term, the unobserved factors leading to such variations are generally left out in the conventional stochastic frontier methods. That is, the pattern of time-variation in efficiency is restricted to follow a known function of exogenous parameters. Hence, one of the main goals of this study is to point out the importance of capturing these time-varying unobserved factors in the efficiency analysis, especially for longer panel data, and the relative ease with which such time-varying unobserved factors can be addressed using the Kalman filter. The results of our Monte Carlo simulations serve well for this purpose. Our model is not unduly complicated and can be applied relatively easily in many applications. Thus the KFE is proposed as a simple and effective (as shown in the simulations) solution to the problem at hand. The KFE can be viewed as an alternative to the factor model approach addressed in Kneip, Sickles, and Song (2012) and Ahn, Lee, and Schmidt (2013) and recent generalizations utilizing Bayesian alternatives.

An early application of the Kalman filter in the productivity setting is Slade (1989) where she uses the local level model with trend to model total factor productivity. However, Ueda and Hoshino (2005) appear to have been the first to apply the Kalman filter to the estimation of efficiency in a data envelopment analysis (DEA) framework. Ueda and Hoshino (2005) examine the case where the inputs and outputs are not deterministic.

3

Kutlu (2010a), Emvalomatis, Stefanou, and Lansink (2011) and our study appear to be the first to use the Kalman filter to estimate efficiency in the framework of stochastic frontier analysis (SFA).[2] Emvalomatis, et al. (2011) modeled the logarithm of ratio of inefficiency and efficiency by a generalized version of an AR(1) process. Their method, however, does not use the traditional Kalman filter since the state variable is not linearly incorporated in their model, which is a necessary assumption for the traditional version of the Kalman filter. Hence, they use a non-linear version of the Kalman filter. In contrast, we model the effects term as in the local level model and calculate the efficiency scores utilizing the approach adopted by Schmidt and Sickles (1984). Moreover, for our model the traditional Kalman filter method is sufficient for our estimation purposes, although extensions of the Kalman filter, for example, to handle endogenous regressors, recently have been developed and used in a production setting.[3] We apply the KFE to estimate the average (and individual) efficiencies of the U.S. airlines during the period 1999-2009. Over our 11 years of study period, the average efficiency of the airlines do not show a tendency to increase. Indeed, for the first few years of the study it seems that the efficiencies of the airlines decreased. As efficiency change and technical (innovation) change are the two main components of productivity growth our empirical findings are broadly consistent with the findings of others (see, for example, Färe et al., 2007) who report declining service quality as problems with delays and congestion at US major airports accelerated during our sample period.

In the next section we describe the KFE and propose several ways in which it can be implemented to model productive efficiency. In section 3 we discuss our Monte Carlo simulation results. Section 4 provides the data description and results of an analysis of

---

[2]Our paper is a substantially revised and extended version of Chapter 2 in Levent Kutlu's dissertation, Market Power and Efficiency (2010a). Recently, independent from us, Peyrache and Rambaldi (2013) proposed a similar Kalman filter model for estimating efficiencies.

[3]For details see Jin and Jorgenson (2011), Kim (2006), Kim and Kim (2011), Kim and Nelson (2006), and Kutlu and Sickles (2012).

productivity trends in the US commercial airline industry during the period 1999-2009. Section 5 concludes. Additional estimation results for other functional specifications as a check of the robustness of our overall findings are provided in the Appendix.

2 Description of the Kalman Filter Estimator

Consider a panel of n_i firms observed over n_t periods. A general stochastic frontier model is given as follows:

$$y_{it} = X_{it}\beta + \mu_{it} + \varepsilon_{it} \tag{1}$$

$$\mu_{i,t+1} = \rho\mu_{i,t} + \tau_{it} + e_{1it}$$

$$\tau_{i,t+1} = \tau_{i,t} + e_{2it}$$

where y_{it} is the logarithm of output, $\varepsilon_{it} \sim \mathbf{NID}\left(0, \sigma_\varepsilon^2\right)$ and $e_{it} = \begin{bmatrix} e_{1it} & e_{2it} \end{bmatrix}' \sim$ $\mathbf{NID}\left(0, Q\right)$ are independently distributed error terms. The initial values of the state variables μ_{it} and τ_{it} are assumed to be jointly normally distributed with zero mean and they are independent from ε_{it} and e_{it}. Estimation details are provided in Appendix A. The component μ_{it} is the random heterogeneity specific to i^{th} individual which is interpreted as efficiency. In the spirit of Ahn, et al. (2000) we allow the firm to sluggishly reduce its inefficiency by modeling efficiency as an AR(1) process with trend τ_{it}. We also allow the firm to adjust quickly. Efficiency may be a random walk (or a random walk with trend), for example (cf, Kneip et al., 2012) and thus the model allows for non-stationarity. In our empirical illustration of the KFE that we explore in section 5, we estimate production efficiency using a restricted version of the translog (RTRANS) production function. The restricted version of the translog that we use provides us with an empirical vehicle that suits our purpose in this introduction of a new estimator and

5

is statistically supported over the full translog model.[4] As a check of the robustness of results based on the restricted translog model we also present estimation results from the full translog model in the Appendix B.

We calculate the time-varying production frontier intercept common to all producers in period t as $\widehat{\mu}_t = \max_i \widehat{\mu}_{it}$ (Cornwell, et al., 1990). Relative technical efficiency is estimated as $TE_{it} = \exp(-\widehat{u}_{it})$, where $\widehat{u}_{it} = \widehat{\mu}_t - \widehat{\mu}_{it}$. Equation system 1 can be rewritten as:

$$
\begin{aligned}
y_{it} &= X_{it}\beta + ZB_{it} + \varepsilon_{it}, & \varepsilon_{it} &\sim \mathbf{NID}\left(0, \sigma_\varepsilon^2\right) & (2) \\
B_{i,t+1} &= TB_{i,t} + e_{it}, & e_{it} &\sim \mathbf{NID}\left(0, Q\right)
\end{aligned}
$$

where

$$
B_{it} = \begin{bmatrix} \mu_{it} \\ \tau_{it} \end{bmatrix}, e_{it} = \begin{bmatrix} e_{1it} \\ e_{2it} \end{bmatrix}, T = \begin{bmatrix} \rho & 1 \\ 0 & 1 \end{bmatrix}, \text{ and } Z = \begin{bmatrix} 1 & 0 \end{bmatrix}.
$$

For the initialization of the Kalman filter, one can use the initial values that are implied by stationarity. In the case of non-stationary states, diffuse priors can be used. One practical choice is setting the mean squared error matrix of the initial states to be a constant multiple of the identity matrix. The constant is chosen by the econometrician and should be a large number. Alternatively, one can utilize an exact diffuse initialization.[5] For the sake of simplicity we prefer using the former diffuse initialization method. The traditional Kalman filter estimation may be numerically unstable due to rounding errors which might cause variances to be non-positive definite during the update process. One solution to this issue is using the square-root Kalman filter. Hence, we further implement

---

[4] In the Kalman filter setting it is possible to estimate a cost function with/without input share equations. For the simultaneous equations setting we do not consider a stochastic frontier model because of so called Greene's problem. See Kumbhakar (1997), Kumbhakar and Lovell (2003), and Kutlu (2013).
[5] See Durbin and Koopman (2001) for more details about initialization.

the square-root Kalman filter.[6]

A simpler and yet flexible model we will use is:

$$y_{it} = X_{it}\beta + \mu_{it} + \varepsilon_{it}, \qquad \varepsilon_{it} \sim \mathbf{NID}\left(0, \sigma_\varepsilon^2\right) \qquad (3)$$

$$\mu_{i,t+1} = \mu_{i,t} + e_{it}, \qquad e_{it} \sim \mathbf{NID}\left(0, Q\right).$$

This model generalizes the panel data models where the effects term is time-invariant by using time-specific local approximations, i.e., Sickles and Schmidt (1984). When $Q = 0$, μ_{it} is a deterministic function of initial values, i.e., $\mu_{it} = \mu_{i0}$. When choosing this model we follow a commonly used modeling of time-varying parameter models (i.e., random walk parameters); and we do not claim that this model is preferred over the more general model presented above. However, the simple model may perform better for relatively shorter panel data applications as diffuse initialization of state variables eats up smaller number of observations.[7] For example, in our empirical model, which uses an unbalanced panel data set with 11 time periods, the full model was not suitable for estimation.[8]

KFE is a random effects-type estimator, in the sense that $E\left[X_{it}\mu_{it}\right] = 0$ is needed for consistency, and is considerably flexible in terms of capturing latent cross-sectional variations that can change over time and which we consider herein unobservable productivity effects. If the ε_{it} or μ_{it} (effects) terms are correlated with the regressors, then the parameter estimates are inconsistent. The KFE can be modified in line with the control function approach used by Kim and Kim (2011) in order to allow for endogenous regressors that are correlated with the ε_{it} term.[9] Kim (2008) provides a solution to a similar

---

[6]See Durbin and Koopman (2001) and Kutlu and Sickles (2012) for details of the square-root Kalman filter.

[7]See Appendix A for more details about required degrees of freedom. The degrees of freedom requirement may be eased by using other (yet restrictive) initialization approaches.

[8]We failed to estimate the full model for this short panel data.

[9]Kutlu (2010b), Karakaplan and Kutlu (2015), and Tran and Tsionas (2012) use similar control function approaches to deal with endogeneity issues in the stochastic frontier context.

endogeneity problem in the context of Markov-switching models when the state variable and regression disturbance are correlated. If the regressors are correlated with the effects term, then we can estimate the first differenced model:

$$\Delta y_{it} \;=\; \Delta X_{it}\beta + e_{it} + \Delta\varepsilon_{it} \tag{4}$$
$$\;=\; \Delta X_{it}\beta + w_{it}$$

by instrumental variables and standard Kalman filter estimation methods can be applied to the consistent residuals, $y_{it} - X_{it}\hat{\beta}$, in order to obtain the consistent hyperparameter estimates.[10]

3 Monte Carlo Experiments

In this section we implement a set of Monte Carlo simulations to examine the finite sample performance of the KFE. For expositional simplicity we consider a production model. The data generating process is given by:[11],[12]

$$y_{it} \;=\; x_{it}\beta + \varepsilon_{it} - \mu_{it}, \;\; \varepsilon_{it} \sim \mathbf{NID}\left(0, \sigma_\varepsilon^2\right) \tag{5}$$
$$x_{i,t+1} \;=\; Rx_{i,t} + \xi_{it}, \;\; \xi_{it} \sim \mathbf{NID}\left(0, I_2\right)$$

---

[10]See Harvey (1989) for more details on this type of solutions to the endogeneity problem in the Kalman filter setting.

[11]When generating regressors we followed Park, Sickles, and Simar (2003, 2007) and Kutlu (2010b).

[12]Note that for the KFE μ_{it} may be negative or positive. Hence, as long as the u_{it} term is predicted properly the sign of μ_{it} is not important. However, in the simulations the production model is written in a general form so that BC production model is also nested. Hence, for this purpose the sign of μ_{it} is negative.

where $x_{it} = \begin{bmatrix} x_{1it} & x_{2it} \end{bmatrix} \sim \mathbf{NID}\left(0, \left(I_2 - R^2\right)^{-1}\right), \beta = \begin{bmatrix} \beta_1 & \beta_2 \end{bmatrix}' = \begin{bmatrix} 0.5 & 0.5 \end{bmatrix}',$
$\sigma_\varepsilon^2 = 1$, and

$$R = \begin{bmatrix} 0.4 & 0.05 \\ 0.05 & 0.4 \end{bmatrix}.$$

The generated values for x are shifted around three different means to obtain three balanced groups of firms. We chose $m_1 = (5, 5)'$, $m_2 = (7.5, 7.5)'$, and $m_3 = (10, 10)'$ as the group means. We simulate a sample of size $(n_i, n_t) = (50, 60)$. Each simulation is carried out $1,000$ times. We consider five different data generating processes for the μ_{it} term:

$$DGP\ 1 \quad : \quad \mu_{it} = \mu_i \tag{6}$$

$$DGP\ 2 \quad : \quad \mu_{it} = a_{0i} + a_{1i}(\frac{t}{n_t}) + a_{2i}(\frac{t}{n_t})^2$$

$$DGP\ 3 \quad : \quad \mu_{it} = b_{0i} + \sum_{r=1}^{2}\{b_{1ri}\sin(\frac{2rt\pi}{n_t}) + b_{2ri}\cos(\frac{2rt\pi}{n_t})\}$$

$$DGP\ 4 \quad : \quad \mu_{it} = \eta_t u_i$$

$$DGP\ 5 \quad : \quad \mu_{it} = r_{it}$$

where $\mu_i \sim \mathbf{NID}\,(0,1)$; $\eta_t = \exp(-h(t - n_t))$, $h = \frac{0.5}{n_t}$, and $u_i \sim \mathbf{NID}^+\,(0,1)$; $a_{li} \sim \mathbf{N}\,(0,1)$; $b_{lri} \sim \mathbf{NID}\,(0,1)$; $r_{i,t+1} = r_{it} + v_{it}$ and $r_{i1} \sim \mathbf{NID}\,(0,1)$; and $v_{it} \sim \mathbf{NID}\,(0,1)$.

We consider five estimators in our simulations. Each of these estimators correspond to one of the DGPs. The estimators are: Fixed effects (FE) estimator, CSS within estimator (CSSW), Fourier estimator (FOE), Battese-Coelli estimator (BC), and KFE. FE, CSSW, and FOE are described as follows:

$$\hat{\beta} = (X'M_Q X)^{-1} X'M_Q y \tag{7}$$

where $M_Q = I - Q(Q'Q)^{-1}Q'$, $Q = diag\,(W_{i\cdot})$ is a block diagonal matrix with $W_{i\cdot}$ matrices

on the diagonal, and $W_{i.}$ is a matrix with rows W_{it}. For example, we have $W_{it} = 1$ for the FE estimator $W_{it} = [1, \frac{t}{n_t}, (\frac{t}{n_t})^2]$ for the CSSW estimator[13], and $W_{it} = [1, \sin(\frac{2t\pi}{n_t}),$ $\sin(\frac{4t\pi}{n_t}), \cos(\frac{2t\pi}{n_t}), \cos(\frac{4t\pi}{n_t})]$ for the FOE.

Excepting the BC estimator, technical efficiencies are estimated as $TE_{it} = \exp(-\widehat{u}_{it})$, where $\widehat{u}_{it} = \max_i \widehat{\mu}_{it} - \widehat{\mu}_{it}$. The BC estimator assumes that $u_{it} = \eta_t u_i$ where $u_i \sim$ $\mathbf{NID}^+ (m, \sigma_u^2)$ and $\eta_t = \exp(-h(t - n_t))$. Let $e_{it} = \varepsilon_{it} - \mu_{it}$. For the BC estimator the efficiency is estimated by:

$$
\begin{aligned}
TE_{it} &= E[\exp(-u_{it})|e_{it}] \\
&= \frac{1 - \Phi(\eta_t \sigma^* - \frac{m_i^*}{\sigma^*})}{1 - \Phi(-\frac{m_i^*}{\sigma^*})} \exp(-\eta_t m_i^* + \frac{1}{2}\eta_t^2 \sigma^{*2})
\end{aligned}
\tag{8}
$$

where $\eta = (\eta_1, \eta_2, ..., \eta_n)'$, Φ represents the distribution function for the normal random variable and

$$
\begin{aligned}
m_i^* &= \frac{m\sigma_\varepsilon^2 - \eta' e_i \sigma_u^2}{\sigma_\varepsilon^2 + \eta'\eta\sigma_u^2} \\
\sigma^{*2} &= \frac{\sigma_u^2 \sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \eta'\eta\sigma_u^2}.
\end{aligned}
$$

For the KFE we assume the following model:

$$
\begin{aligned}
y_{it} &= X_{it}\beta + \mu_{it} + \varepsilon_{it}, & \varepsilon_{it} &\sim \mathbf{NID}\left(0, \sigma_\varepsilon^2\right) \\
\mu_{i,t+1} &= \mu_{i,t} + e_{it}, & e_{it} &\sim \mathbf{NID}\left(0, Q\right).
\end{aligned}
\tag{9}
$$

Hence, for the KFE the effects term is modelled as a random walk, which is consistent with the local level model of univariate time series. We provide the bias, the variance, the mean squared error (MSE) of the coefficients, the (normalized) MSE of the efficiency

---

[13]The original CSSW estimator assumes $W_{it} = [1, t, t^2]$. However, for the simulations we normalize t by n_t. This normalization does not affect the results and is done for numerical purposes.

estimates as well as the Pearson and Spearman correlations of efficiency estimates with the true efficiency levels. The MSE of the efficiencies are calculated as follows:

$$MSEeff(TE_{0it}, \widehat{TE}_{it}) = \frac{\sum_{i,t}\left(TE_{0it} - \widehat{TE}_{it}\right)^2}{\sum_{i,t} TE_{0it}^2} \qquad (10)$$

where TE_{0it} is the true technical efficiency level and $\widehat{TE}_{it}$ is the estimated efficiency level. The results for the Monte Carlo experiments are given in Table 1-5.

Table 1. Monte Carlo Results for DGP1 (FE)

	FE	CSSW	FOE	BC	KFE
MSE	0.0006	0.0007	0.0007	0.0008	0.0006
Bias1	0.0002	0.0002	0.0004	-0.0046	0.0002
Bias2	0.0000	-0.0002	-0.0004	-0.0049	0.0000
Var1	0.0003	0.0003	0.0003	0.0004	0.0003
Var2	0.0003	0.0003	0.0004	0.0004	0.0003
MSEeff	0.0180	0.0528	0.0873	0.1109	0.0720
CORP	0.9999	0.9995	0.9991	0.9933	0.9983
CORS	1.0000	0.9989	0.9994	0.9987	0.9978

Table 2. Monte Carlo Results for DGP2 (CSSW)

	FE	CSSW	FOE	BC	KFE
MSE	0.0007	0.0006	0.0007	0.0009	0.0006
Bias1	-0.0009	-0.0008	-0.0010	-0.0077	-0.0008
Bias2	-0.0007	-0.0003	-0.0006	-0.0072	-0.0005
Var1	0.0004	0.0003	0.0003	0.0004	0.0003
Var2	0.0004	0.0003	0.0004	0.0004	0.0003
MSEeff	0.0606	0.0413	0.0776	0.1599	0.0955
CORP	0.9679	0.9985	0.9905	0.9621	0.9926
CORS	0.8882	0.9989	0.9740	0.8623	0.9877

Table 3. Monte Carlo Results for DGP3 (FOE)

	FE	CSSW	FOE	BC	KFE
MSE	0.0031	0.0015	0.0007	0.0014	0.0008
Bias1	0.0010	0.0006	0.0001	0.0043	0.0006
Bias2	0.0004	-0.0002	0.0001	0.0048	-0.0000
Var1	0.0016	0.0007	0.0003	0.0007	0.0004
Var2	0.0016	0.0007	0.0003	0.0007	0.0004
MSEeff	3.2996	0.7278	0.1332	5.6036	0.3621
CORP	0.0547	0.3405	0.9705	0.1711	0.8657
CORS	0.0152	0.5906	0.9986	0.0599	0.9695

Table 4. Monte Carlo Results for DGP4 (BC)

	FE	CSSW	FOE	BC	KFE
MSE	0.0006	0.0007	0.0007	0.0005	0.0006
Bias1	0.0002	0.0002	0.0004	-0.0050	0.0002
Bias2	-0.0000	-0.0002	-0.0004	-0.0052	-0.0001
Var1	0.0003	0.0003	0.0003	0.0002	0.0003
Var2	0.0003	0.0003	0.0004	0.0002	0.0003
MSEeff	0.0352	0.0845	0.1336	0.0203	0.0993
CORP	0.9985	0.9858	0.9991	0.9890	0.9470
CORS	0.9963	0.9826	0.9980	0.9981	0.9421

Table 5. Monte Carlo Results for DGP5 (KFE)

	FE	CSSW	FOE	BC	KFE
MSE	0.0151	0.0041	0.0046	0.0116	0.0014
Bias1	0.0006	-0.0003	0.0007	0.0351	0.0002
Bias2	0.0003	-0.0007	0.0002	0.0332	-0.0013
Var1	0.0077	0.0021	0.0023	0.0048	0.0007
Var2	0.0074	0.0020	0.0023	0.0045	0.0007
MSEeff	1.0813	0.5592	0.4246	1.3638	0.1856
CORP	0.5032	0.7408	0.7959	0.5288	0.9713
CORS	0.5644	0.9634	0.9214	0.7201	0.9975

For the β estimates, the estimators generally show similar performances. For both the β estimates and the efficiency estimates we find that whenever there is a high variation in the efficiency term the less flexible estimators, such as FE and BC, perform worse than the others. KFE performs particularly well in terms of correlations between the true efficiency and the estimated efficiency. It is worth noting that all estimators other than the FOE

and the KFE performed very poorly for DGP3. Indeed, the FE and BC estimators show almost no correlation between the true efficiency and the estimated efficiency.[14] This is because these estimators are not flexible enough to capture the time-varying pattern of the efficiency. Hence, this simulation study shows that when the efficiencies of the firms fluctuate the performance of non-flexible efficiency estimators can be arbitrarily misleading in capturing the performances of firms.

Finally, we present simulation results for smaller sample sizes in Table 6-7: $(n_i, n_t) = (50, 10)$ and $(n_i, n_t) = (10, 60)$. As in our earlier simulations the estimators performed more or less the same in terms of β estimates. Hence, we only summarize their performance for efficiency estimation. The last two rows are the averages of MSE values and Pearson correlations which may serve as an aggregate measure of performance. These tables also confirm that the KFE estimator performs quite well in terms of capturing the unobserved efficiency. A striking observation is that KFE performs well even for relatively shorter panels.

---

[14]In some of the simulation runs we observed even negative correlations.

Table 6. Monte Carlo Results for n_i=50, n_t=10

		FE	CSSW	FOE	BC	KFE
DGP1 (FE)	MSE	0.1169	0.3062	0.4746	0.4955	0.2684
	CORP	0.9987	0.9933	0.9909	0.9879	0.9917
DGP2 (CSSW)	MSE	0.1300	0.2311	0.3902	0.9571	0.2784
	CORP	0.9663	0.9923	0.9838	0.9609	0.9827
DGP3 (FOE)	MSE	1.4859	0.7420	0.5297	4.4685	0.5754
	CORP	0.0149	0.3819	0.7725	0.1623	0.7157
DGP4 (BC)	MSE	0.1682	0.3355	0.4668	0.0741	0.2969
	CORP	0.9982	0.9359	0.9972	0.9485	0.8970
DGP5 (KFE)	MSE	0.6262	0.2775	0.3741	1.1006	0.3429
	CORP	0.6814	0.9363	0.9313	0.5869	0.9566
Aggr.	MSE	0.5054	0.3785	0.4471	1.4192	0.3524
	CORP	0.7319	0.8479	0.9351	0.7293	0.9087

Table 7. Monte Carlo Results for n_i=10, n_t=60

		FE	CSSW	FOE	BC	KFE
DGP1 (FE)	MSE	0.0262	0.0561	0.0840	0.0654	0.0671
	CORP	0.9999	0.9977	0.9996	0.9947	0.9841
DGP2 (CSSW)	MSE	0.0914	0.0538	0.0961	0.8123	0.1048
	CORP	0.8801	0.9945	0.9692	0.8806	0.9589
DGP3 (FOE)	MSE	0.4887	0.2218	0.1007	0.6600	0.1999
	CORP	0.0221	0.6431	0.9797	0.3130	0.8871
DGP4 (BC)	MSE	0.0308	0.0620	0.0914	0.2489	0.0710
	CORP	0.9968	0.9875	0.9978	0.9918	0.9479
DGP5 (KFE)	MSE	0.6339	0.4204	0.3331	0.9578	0.1432
	CORP	0.6337	0.7673	0.8253	0.6533	0.9723
Aggr.	MSE	0.2542	0.1628	0.1411	0.5489	0.1172
	CORP	0.7065	0.8780	0.9543	0.7667	0.9501

4 The U.S. Airline Industry 1999-2009

4.1 The Data

In order to illustrate our estimator and its usefulness in applied settings, we utilize annual data from the U.S. airline industry during the period 1999-2009. The third author has written extensively on commercial airline efficiency issues in the U.S., Europe, and in Asia. We view the example below as informative in regard to the usefulness of our estimator in modeling efficiencies in the airline industry and how it may inform researchers in more extensive industry studies as to potential limitations in their modeling approaches and alternative approaches they may wish to consider, such as ours. The time period we choose is one during which the U.S. airlines faced serious financial troubles. The financial

losses for domestic passenger airline operations were more than three times the losses between 1979-1999. Some of the exogenous cost shocks during the sample period were due to increased taxes and jet fuel prices. At the same time fares fell and remained relatively low. Real jet fuel prices were about 20% lower in 2009 than in 2000. Since 1979 demand grew steadily. However, we observe sharp demand drops during the recession of 2001-2002 and 2008-2009. Due to capital costs and sticky labor prices such unanticipated decreases in demand brought additional complications to an industry which had been experiencing relatively stable and steady demand growth. Another feature of the sample time period is the increase in load factors. Average load factors increased from 71% to 81% between 2000 and 2009 due in part to improved yield management techniques and reduced flight frequency but which also lead to reduced levels of service quality.[15]

The unbalanced data is mainly obtained from the International Civil Aviation Organization (ICAO). The data set that we use has 35 airlines and 298 observations.[16] Input and output variables are constructed following the approaches of Sickles (1985) and Sickles et al. (1986). Inputs are flight capital (K, quantity of planes), labor (L, quantity of pilots, cabin crew, mechanics, passenger and aircraft handlers, and other labour), fuel (F, quantity of barrels of fuel), and materials (M, quantity index of supplies, outside services, and non-flight equipment's). We focus on value added from capital and labor in our empirical illustration of the KFE by netting out from revenue output (RTK, revenue ton kilometers) the value of the intermediate energy and materials. Thus our technology is rather simple and uses capital and labor to produce value added revenue ton kilometers.

In addition to the above, we include two sets of control variables into our model to account for the heterogeneity of output and the capital input. The first set of control variables is concerned with service characteristics: (i) aircraft stage length (SL) and (ii)

---

[15]For more information about the financial situations of U.S. airlines see Borenstein (2011).

[16]The full data set has 39 airlines and 321 observations. We droped 1 airline with less than 4 observations and 3 cargo airlines.

load factor (LF). SL is the average length of a route segment, obtained by dividing the miles flown by the number of departures. The shorter (low value) the stage length the shorter the period an airlines' aircraft spends in each flight segment. LF reflects the average occupancy of an airline's aircraft seats, is considered a measure of service quality, and is often used as a proxy for service competition. A lower load factor often implies that the airline assigns a relatively larger number of planes to a particular route and reflects higher service quality by the airline. The second set of control variables is concerned with capital stock characteristics. The first is the average size of the airline's aircraft (SIZE). The larger the size of the aircraft the more services can be provided without a proportionate increase in factors such as flight crew, passenger and aircraft handlers, and landing slots. The second is the percentage of each airline fleet that is a (JET) aircraft to total number of aircraft. JET is considered as a proxy for the aircraft speed. The jet aircraft tends to fly around three times as fast as turboprops aircraft and in addition the jet aircraft requires a relatively lower number of flight crew resources. A brief description of the variables is given in Table 8.

Table 8: Description of Variables

Variable	Description	Min	%25 Perc.	%75 Perc.	Max	Mean	Std
Q (y)	ln(Value added RTK)	10.9126	12.6098	15.0266	16.8962	13.9445	1.6004
QL (x_1)	ln(Labor quantity)	3.7351	6.5181	8.5467	10.2650	7.4982	1.5112
QK (x_2)	ln(Capital quantity)	1.9459	3.4410	5.6113	6.7038	4.5095	1.2414
LF (x_3)	Load factor	0.1500	0.5300	0.6210	0.8030	0.5731	0.0990
SL (x_4)	ln(Stage length)	5.5968	6.5164	7.5368	8.5643	7.0730	0.6636
JET (x_5)	Jet engines	0.0000	0.9003	1.0000	1.0000	0.8722	0.2614
$SIZE$ (x_6)	ln(Average plane size)	2.7568	4.0943	5.2244	5.8926	4.7579	0.6020

4.2 Analysis using the KFE

In this section we examine the technical efficiency trends in the U.S. airline industry during the period 1999-2009 using our new KFE and compare our findings to those from the Battese-Coelli (BC) and the Cornwell, Schmidt, and Sickles within (CSSW) estimator with efficiency modeled as depending only on deterministic time period proxies that vary over time. We utilize the quadratic specification used in the U. S. airline empirical illustration of Cornwell, Schmidt, and Sickles (1990). The BC estimator is probably the most widely used of the panel estimators and is a random effects type estimator of efficiency change that also utilizes a deterministic time trend. The CSSW has somewhat more flexibility and provides a fixed effects treatment. We estimate the value-added production function of the U.S. airlines (revenue ton kilometers less a value weighted average of materials and energy). The production function is specified as linear in logs as:

$$y_{it} = \sum_j \beta_j x_{jit} + \mu_{it} + \varepsilon_{it} \tag{11}$$

$$\mu_{i,t+1} = \mu_{i,t} + e_{it}$$

where $\varepsilon_{it} \sim \mathbf{NID}(0, \sigma_\varepsilon^2)$ and $e_{it} \sim \mathbf{NID}(0, \sigma_e^2)$ are independently distributed error terms.[17] The estimates for the production function parameters and average efficiencies for the KFE and the BC and CSSW are given in Table 9 and Figure 1, respectively. [18] The overall average efficiencies for the KFE and the CSSW and BC estimators are 0.577, 0.438, and 0.632, respectively.

The median of the returns to scale values for the KFE and the BC and CSSW estima-

---

[17]As mentioned in our theoretical section we want to concentrate on the simple local level model rather than the full model as it is easier to estimate. For example, our attempt to estimate the full model failed. The alternative estimates for random walk with deterministic trend and AR(1) effects models are provided in Appendix B.

[18]When calculating the efficiency estimates, we trim the effects term from the upper and lower 7.5% percentiles, observed at least at one time period, to remove the outlier effects. See, Berger (1993), Berger and Hannan (1998), Kutlu (2012), and Kneip, Sickles, and Song (2012) for more details. See, also Appendix B for some robustness check for trimming.

tors are 0.883, 0.94, and 1.034, respectively. A common finding for the airline industry is that the airlines operate in a constant returns to scale environment. In a single-output production setting, Basu and Fernald (1997) provide a theoretical proof that the value added estimate of returns to scale is smaller (greater) than the corresponding gross output model when there is decreasing (increasing) returns to scale. Hence, there is a magnification effect for returns to scale estimates when a value added production function is used.[19] Therefore, the returns to scale estimate for the KFE might have been driven by this fact. Nevertheless, the constant returns to scale value of 1 lies within one sample standard deviation away from the mean value of returns to scale estimates from the KFE. In terms of regularity conditions, KFE outperforms other two estimators. More precisely, while the KFE satisfies curvature regularity condition at each time period, the BC and CSSW violate curvature regularity condition at each time period. At the median values of the regressors, all three estimators satisfy monotonicity conditions at each time period. According to KFE estimates, the average efficiency of the U.S. airlines is relatively stable for the second half of the study period. However, there is some evidence of a decrease in efficiency for the first half of the study period.

One potential empirical concern would be whether the effects term is correlated with the regressors or not. If the effects term is correlated with the regressors, then the coefficient estimates would be inconsistent. One advantage of the CSSW estimator over the random effects-type estimator is that even when the regressors are correlated with the effects term, the parameter estimates are consistent. Hence, the parameter estimates from the CSSW model can be used to a test the consistency of parameter estimates from the KFE. We test the consistency of parameter estimates from KFE using a Wu-Hausman test and cannot reject the KFE estimates at the 5% significance level.

---

[19]For similar results see also Diewert and Fox (2008).

Table 9. Restricted Translog Production Model Estimates

	KFE			CSSW			BC		
	Param.	SE	P-val	Param.	SE	P-val	Param.	SE	P-val
Q_L	0.5413 *	0.3269	0.098	0.9000 **	0.3963	0.024	0.4320	0.2980	0.147
Q_K	1.1789 **	0.4748	0.013	2.1244 **	0.9209	0.022	1.1235 ***	0.3463	0.001
LF	-7.5881 ***	2.4679	0.002	-5.4241	3.9615	0.173	-7.6096 ***	2.2275	0.001
SL	3.2997 **	1.4505	0.023	6.4690 **	2.9879	0.032	1.5974 *	0.8818	0.070
JET	0.8529	3.5434	0.810	-1.6506	6.1859	0.790	3.4185	2.4065	0.155
SIZE	1.7849 **	0.8042	0.027	0.4726	1.2492	0.706	1.4273 **	0.7143	0.046
$Q_L^2/2$	0.0387	0.1071	0.717	0.1126	0.1415	0.427	0.1560	0.0958	0.103
$Q_L Q_K$	-0.1550	0.1417	0.274	-0.3351	0.2088	0.110	-0.2896 **	0.1196	0.015
$Q_K^2/2$	0.1540	0.1994	0.440	0.2232	0.2603	0.392	0.3910 **	0.1729	0.024
$LF^2/2$	1.3352	1.7594	0.448	2.0109	2.1742	0.356	0.4959	1.7470	0.777
LF*SL	0.2704	0.7087	0.703	0.5637	1.0427	0.589	0.2490	0.6580	0.705
LF*JET	-1.3300	1.1226	0.236	-0.8783	1.5403	0.569	-2.0094 *	1.0592	0.058
LF*SIZE	1.4470 *	0.7504	0.054	0.3446	0.9734	0.724	1.7017 **	0.6900	0.014
$SL^2/2$	-0.2324	0.2446	0.342	-1.4053 **	0.5994	0.020	0.2561	0.1722	0.137
SL*JET	0.2379	0.6702	0.723	0.7894	1.2296	0.522	-0.4244	0.4849	0.381
SL*SIZE	-0.3333	0.2552	0.192	0.5557	0.4183	0.186	-0.4800 ***	0.1792	0.007
$JET^2/2$	-0.1142	1.4846	0.939	-1.0271	2.4465	0.675	0.9249	1.2021	0.442
JET*SIZE	-0.2003	0.3192	0.5304	-0.3082	0.4862	0.5270	0.0541	0.2575	0.8340
$SIZE^2/2$	0.0252	0.3280	0.9388	-0.9068 *	0.5226	0.0840	0.2835	0.2214	0.2000
σ_ε	0.1263 ***	0.0150	0.0000	-	-	-	-	-	-
σ_e	0.1124 ***	0.0203	0.0000	-	-	-	-	-	-
Constant	-	-	-	-	-	-	-5.2215	3.2784	0.1110

Note: P-val<0.10: * , P-val<0.05: **, and P-val<0.01: ***

Figure 1: Efficiency Estimates for Restricted Translog Model
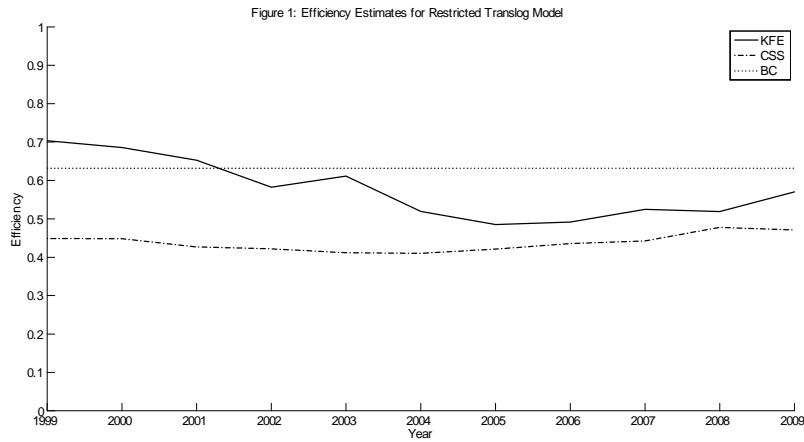
We also check the robustness of our results by estimating a full version of the translog model. A common problem with the translog production function is that by increasing the number of variables by adding second-order ln terms to the Cob-Douglas functional form the second order terms tend to exhibit considerable multicollinearity. The full translog model estimates are given in the Appendix. For the full translog model, few of the

parameters were significant at the 5% level. We choose our final model specification based on the BIC for the Kalman filter. This criterion is:

$$BIC = \frac{-2\ln L + \ln(s)(p+d)}{s}$$

where L is the likelihood value, s is the sample size, p is the number of hyperparameters, and d is the number of diffuse priors (Durbin and Koopman, 2001). The BIC values for the full translog and restricted translog forms are 1.805 and 1.714, respectively. Based on the BIC and the fact that almost all the parameters of the full translog model are insignificant, the restricted translog functional form is preferred on statistical grounds.

5   Conclusions

In this study we have proposed a way to measure technical efficiency via the Kalman filter. Our new Kalman Filter estimator (KFE) provides a local approximation to general time and cross sectionally varying effects terms in a standard panel model. We examine the new estimator in a series of limited Monte Carlo experiments. Our simulation results indicate that while the performance of the KFE is similar to the performances of the other estimators for the coefficient estimates, the KFE outperforms the less flexible estimators in terms of the correlation of the effects with true effects. A result of our simulations is that the widely used BC estimator performed very poorly whenever there is substantial variation in the effects, or for our canonical stochastic frontier efficiency model, the efficiency term. If the sample data contains events that can cause jumps in the productivity of firms, then the KFE estimator appears able to improve on other standard panel treatments that are less flexible in specifying the temporal variation in the effects. We then used the KFE in order to estimate the average efficiency of the U.S.

airlines. Point estimates for the KFE indicate that average efficiency of the U.S. airlines fell by more than 10% during earlier years of time period, but these trends are not stable. What does appear to be the case is that there is no strong or even weak evidence that airlines experienced improved efficiencies over the sample period. Given that there were no particularly important new technical innovations during the sample period, the sizeable losses incurred by the industry as fares continued to be held down by competitive pressures were not surprising. Moreover, many of the efficiency gains from reorganized networks, code-sharing arrangements, and other best business practices apparently had already been realized by the beginning of the sample period.

6   References

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

Ahn, S. C., Lee, Y. H., and Schmidt, P. (2013). Panel Data Models with Multiple Time-varying Individual Effects. *Journal of Econometrics,* 174, 1-14.

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

Assaf, A.G., Gillen, D., and Tsionas, E.G. (2014), Understanding Relative Efficiency Among Airports: A General Dynamic Model for Distinguishing Technical and Allocative Efficiency, Transportation Research Part B, 70, 18-34.

Battese, G.E. and Coelli, T.J. (1992), Frontier Production Functions, Technical Efficiency and Panel Data with Application to Paddy Farmers in India, *Journal of Productivity Analysis*, 3, 153-169.

Basu, S. and Fernald, J.G. (1997), Returns to Scale in U.S. Production: Estimates and Implications, Journal of Political Economy, 105, 249-83.

Berger, A.N. (1993), Distribution Free Estimates of Efficiency in US Banking Industry and Tests of the Standard Distributional Assumptions, *Journal of Productivity Analysis*, 4, 261-292.

Berger, A.N. and Hannan, T.H. (1998), The Efficiency Cost of Market Power in the Banking Industry: A Test of the "Quiet Life" and Related Hypotheses, *Review of Economics and Statistics*, 80, 454-465.

Borenstein, S. (2011), On the Persistent Financial Losses of U.S. Airlines: A Preliminary Exploration, Working paper.

Cornwell, C., Schmidt, P., and Sickles, R.C. (1990), Production Frontiers with Time-Series Variation in Efficiency Levels, *Journal of Econometrics* 46, 185-200.

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

Diewert, W.E. and Fox, K.J (2008), On the Estimation of Returns to Scale, Technical Progress and Monopolistic Markups, Journal of Econometrics, 145, 174-193.

Durbin, J. and Koopman, S.J. (2001), Time Series Analysis by State Space Methods, Oxford Statistical Series 24, Oxford University Press.

Emvalomatis, G., Stefanou, S.E., and Lansink A.O. (2011), A Reduced-Form Model for Dynamic Efficiency Measurement: Application to Dairy Farms in Germany and The Netherlands, *American Journal of Agricultural Economics*, 93, 161-174.

Färe, R., Grosskopf, S., and Sickles, R.C. (2007), Productivity? of U.S. Airlines After Deregulation, *Journal of Transport Economics and Policy*, 41, 93-112.

Galán, J.E. and Pollitt, M.G. (2014), Inefficiency Persistence and Heterogeneity in Colombian Electricity Utilities, Energy Economics, 46, 31-44.

Harvey, A.C (1989), Forecasting, Structural Time Series Models and the Kalman Filter, Cambridge University Press, Cambridge.

Huang, T.H. and Chen, Y.H. (2009), A Study on Long-Run Inefficiency Levels of a Panel Dynamic Cost Frontier under the Framework of Forward-Looking Rational Expectations, *Journal of Banking and Finance*, 33, 842-849.

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

Jin, H. and Jorgenson, D.W. (2009), Econometric Modeling of Technical Change, *Journal of Econometrics*, 157, 205-219.

Kalman, R.E. (1960), A New Approach to Linear Filtering and Prediction Problems, *J. Basic Engineering, Transactions ASMA*, Series D, 82, 35-45.

Karakaplan, M.U. and Kutlu, L. (2015), Consolidation Policies and Saving Reversals,

Working paper (SSRN: http://ssrn.com/abstract=2607276).

Kim, C.J. (2006), Time-Varying Parameter Models with Endogenous Regressors, *Economics Letters*, 91, 21-26.

Kim, C.J. (2008), Dealing with Endogeneity in Regression Models with Dynamic Coefficients, Foundations and Trends in Econometrics, 3, 165-266.

Kim, Y. and Kim C.J. (2011), Dealing with Endogeneity in a Time-Varying Parameter Model: Joint Estimation and Two-Step Estimation Procedures, *Econometrics Journal*, 14, 487-497.

Kim, C.J. and Nelson , C.R. (2006), Estimation of a Forward-Looking Monetary Policy Rule: A Time-Varying Parameter Model Using Ex Post Data, *Journal of Monetary Economics*, 53, 1949-1966.

Kumbhakar, S.C. (1990), Production Frontiers, Panel Data, and Time-Varying Technical Inefficiency, *Journal of Econometrics*, 46, 201-211.

Kumbhakar, S.C. (1997), Modeling Allocative Inefficiency in a Translog Cost Function and Cost Share Equations: An Exact Relationship, *Journal of Econometrics*, 76, 351-356.

Kumbhakar, S.C. and Lovell, C.A.K. (2003), Stochastic Frontier Analysis, Cambridge University Press.

Kutlu, L. (2010 b), Battese-Coelli Estimator with Endogenous Regressors, *Economics Letters*, 109, 79-81.

Kutlu, L. (2010 a), Market Power and Efficiency, Unpublished PhD dissertation, Rice University.

Kutlu, L. (2012), U.S. banking efficiency, 1984-1995, *Economics Letters*, 117, 53-56.

Kutlu, L. (2013), Misspecification in Allocative Inefficiency: A Simulation Study, *Economics Letters*, 118, 151-154.

Kutlu, L. and Sickles, R.C. (2012), Estimation of Market Power in the Presence of Firm Level Inefficiencies, *Journal of Econometrics*, 168, 141-155.

Kneip, A., Sickles, R.C., and Song, W. (2012), A New Panel Data Treatment for Heterogeneity in Time Trends, *Econometric Theory*, 28, 560-628.

Meeusen, W. and van den Broeck, J. (1977), Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error, *International Economic Review*, 2, 435-444.

Park, B.U., Sickles, R.C., and Simar, L. (2003), Semiparametric Efficient Estimation of AR1 Panel Data Models, *Journal of Econometrics,* 117, 279-309.

Park, B.U., Sickles, R.C., and Simar, L. (2007), Semiparametric Efficient Estimation of Dynamic Panel Data Models, *Journal of Econometrics*, 136, 281-301.

Peyrache, A. and Rambaldi, A.N. (2013), Incorporating Temporal and Country Heterogeneity in Growth Accounting - An Application to EU-KLEMS, Working paper.

Pitt, M.M. and Lee, L.F. (1981), The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry, Journal of Development Economics, 9, 43-64.

Schmidt, P., and Sickles, R.C. (1984), Production Frontiers and Panel Data, *Journal of Business and Economic Statistics*, 2, 367-374.

Sickles, R.C. (1985), A Nonlinear Multivariate Error-Components Analysis of Technology and Specific Factor Productivity Growth with an Application to the U.S. Airlines, *Journal of Econometrics,* 27, 61-78.

Sickles, R.C. (2005), Panel Estimators and the Identification of Firm-Specific Efficiency Levels in Parametric, Semiparametric, and Nonparametric Settings, *Journal of Econometrics*, 126, 305-334.

Sickles, R.C., D.H. Good, and Johnson, R.L. (1986), Allocative Distortions and the Regulatory Transition of the U.S. Airline Industry, *Journal of Econometrics,* 33, 143-163.

Sickles, R. C., J. Hao, and Sheng, C. (2015), Panel Data and Productivity Measurement, Chapter 17 *Oxford Handbook of Panel Data*, Badi Baltagi (ed.), New York: Oxford University Press, 517-547.

Slade, M.E. (1989), Modeling Stochastic and Cyclical Components of Technical Change:

An Application of the Kalman Filter, *Journal of Econometrics*, 41, 363-383.

Tran, K.C. and Tsionas, E.G. (2012), GMM Estimation of Stochastic Frontier Model with Endogenous Regressors, Economics Letters, 118, 233-236.

Tsionas, E.G. (2006), Inference in Dynamic Stochastic Frontier Models, *Journal of Applied Econometrics*, 21, 669-676.

Ueda, T. and Hoshino, K. (2005), Estimation of Firms Efficiencies Using Kalman Filter and Stochastic Efficiency Model, *Journal of the Operations Research Society of Japan*, 308-317.

7 Appendix A

In this appendix we provide further details about the Kalman filter estimation. Consider

two stochastic frontier models that are nested by the general setting that we described:

$$
y_{it} = X_{it}\beta + ZB_{it} + \varepsilon_{it}, \qquad \varepsilon_{it} \sim \mathbf{NID}\left(0, \sigma_\varepsilon^2\right) \tag{12}
$$

$$
B_{i,t+1} = TB_{i,t} + Re_{it}, \qquad e_{it} \sim \mathbf{NID}\left(0, \sigma_e^2\right)
$$

where

$$
B_{it} = \begin{bmatrix} \mu_{it} \\ \tau_{it} \end{bmatrix}, T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \ R = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \text{ and } Z = \begin{bmatrix} 1 & 0 \end{bmatrix}.
$$

and

$$
y_{it} = X_{it}\beta + ZB_{it} + \varepsilon_{it}, \qquad \varepsilon_{it} \sim \mathbf{NID}\left(0, \sigma_\varepsilon^2\right) \tag{13}
$$

$$
B_{i,t+1} = TB_{i,t} + e_{it}, \qquad e_{it} \sim \mathbf{NID}\left(0, \sigma_e^2\right)
$$

where

$$
B_{it} = \mu_{it}, \ T = \rho, \text{ and } Z = 1.
$$

The first model assumes a random walk with deterministic trend effects term and the

second model assumes a potentially non-stationary AR(1) process for the effects term.

Both of these models can be estimated by using the recursive equations provided below.

The estimation consists of two steps. Kalman filtering and smoothing. In the first step,

the following recursive Kalman filter equations are applied:

$$\eta_{it} = y_{it} - X_{it}\beta - Zb_{it} \tag{14}$$

$$F_{it} = ZP_{it|t-1}Z' + \sigma_{\varepsilon}^2$$

$$M_{it} = P_{it|t-1}Z'$$

$$b_{it|t} = b_{it|t-1} + M_{it}F_{it}^{-1}\eta_{it}$$

$$P_{it|t} = P_{it|t-1} - M_{it}F_{it}^{-1}M_{it}'$$

$$b_{it|t-1} = Tb_{i,t-1|t-1}$$

$$P_{it|t-1} = TP_{i,t-1|t-1}T' + \sigma_e^2.$$

In the second step, the smoothing is applied by using the following recursive equations:

$$L_{it} = T - TM_{it}F_{it}^{-1}Z \tag{15}$$

$$r_{i,t-1} = Z'F_{it}^{-1}\eta_{it} + L_{it}'r_{it}$$

$$N_{i,t-1} = Z'F_{it}^{-1}Z + L_{it}'N_{it}L_{it}$$

$$\tilde{b}_{it|t-1} = b_{it|t-1} + P_{it|t-1}r_{i,t-1}$$

$$V_{it} = P_{it|t-1} - P_{it|t-1}N_{i,t-1}P_{it|t-1}$$

where $r_{in_t} = 0$ and $N_{in_t} = 0$.[20] The log-likelihood is given by:

$$\ln(L) = \sum_{i=1}^{n_i} L_i = constant - \frac{1}{2}\sum_{i=1}^{n_i}\sum_{t=d_i+1}^{n_t}\left(\ln(F_{it}) + \frac{\eta_{it}^2}{F_{it}}\right) \tag{16}$$

---

[20]Note that smoothing is not needed to get the MLE estimates. The smoothing equations are calculated after the estimations. The Kalman filter uses past and current observations to predict the state variables; and thus it does not use all information when calculating the state variable predictions. Once the parameters of the model are estimated by MLE, the smoothing enables us to update our predictions using information from all time periods. This is why after smoothing procedure the predictions of state variables look smoother. Hence, the name smoothing.

where d_i is the number of diffuse states for firm i. The number of diffuse priors (per panel unit) for the first model is two. The number of diffuse priors (per panel unit) for the second model is one. If we assume that $\rho = 1$, the second model would still have one diffuse prior per panel unit. However, obviously, the number of parameters to be estimated would be smaller. If we assume that $|\rho| < 1$, the second model would not have any diffuse priors.

Let m be the number of state variables (per panel unit) and q be the number of state variables with diffuse priors (per panel unit). For diffuse initialization we assume that:

$$b_{i0} = A\delta + S\eta_{i0}$$

$$\eta_{i0} \sim \mathbf{NID}(0, Q_0)$$

where δ is a $q \times 1$ vector of unknown quantities and the $m \times q$ matrix A and $m \times (m-q)$ matrix S are selection matrices that consist of columns of identity matrix. Then, matrix for initialization is:

$$P_{i0} = \kappa P_{i\infty 0} + P_{*0}$$

where $\kappa \to \infty$, $P_\infty = A'A$, and $P_* = SQ_0S'$. As it can be seen from the log-likelihood, the first d_i observation(s) for panel unit i are burnt out for the sake of initialization and are not considered in the log-likelihood. Hence, for example, for the second model the first observation of each panel unit is used for initialization. The reason for this is that as long as $t \geq d_i + 1$, we would have $P_{i\infty t} = 0$. The variance matrix can be estimated using the standard maximum likelihood procedures. For the estimations we used the standard BFGS optimization method.

8 Appendix B

In this appendix we present additional results based on the full translog model and our truncation scheme when calculating the efficiency estimates for KFE and CSSW estimator. We also provide estimates for alternative Kalman filter models.

The full translog estimates are given in Table 10. The parameter estimates are generally not significant even at 10% significance level. The median of the returns to scale values for the KFE, CSSW, and BC estimators are 0.8625, 1.1478, and 1.0184, respectively. The corresponding returns to scale estimates from the restricted model were 0.883, 0.94, and 1.034, respectively. Hence, for the KFE and BC estimator the returns to scale estimates are robust to the choice of the functional form. Nevertheless, for both restricted and unrestricted translog production models the constant returns to scale value of 1 lies within one sample standard deviation away from the median value of returns to scale estimates from each of these estimates. All the estimators satisfy the monotonicity conditions at the median values of the regressors at each time period. In contrast to the restricted translog production model where only KFE satisfied the regularity conditions at the median values of the regressors, KFE and CSSW estimator satisfies the curvature conditions at each time period. BC estimator failed to satisfy the regularity conditions at four of the time periods. The estimates for the production function parameters and average efficiencies for the KFE and the BC and CSSW estimators are given in Table 10 and Figure 2. The overall average efficiencies for the KFE, CSSW, and BC are 0.637, 0.458, and 0.605, respectively. These values are not substantially different from their restricted counterparts, i.e., 0.577, 0.438, and 0.632. The average efficiencies for the full translog model are provided in Figure 2. In line with the restricted translog model, KFE predicts decrease in efficiency in first few years of the study period and relatively stable efficiency levels for the last couple of years.

Table 10. Full Translog Production Model Estimates

	KFE			CSSW			BC		
	Param.	SE	P-val	Param.	SE	P-val	Param.	SE	P-val
Q_L	-0.0089	1.0890	0.994	0.7998	1.5907	0.616	0.0014	0.9090	0.999
Q_K	-0.1172	1.2252	0.924	-1.3316	2.2313	0.551	1.0972	1.0054	0.275
LF	-5.7379	3.5219	0.103	-10.0304 **	4.8566	0.040	-3.3730	3.2929	0.306
SL	1.5702	1.6977	0.355	3.0570	3.9807	0.444	0.6302	1.0701	0.556
JET	1.4254	3.7352	0.703	-0.7014	7.8046	0.929	2.7640	2.6859	0.303
SIZE	2.0494 *	1.1825	0.083	2.5571	1.7101	0.137	1.4932	1.0830	0.168
$Q_L^2/2$	0.0246	0.1242	0.843	0.0402	0.1803	0.824	0.1643	0.1155	0.155
$Q_L Q_K$	-0.0531	0.1672	0.751	-0.1901	0.2575	0.461	-0.2372	0.1499	0.113
$Q_K^2/2$	0.0446	0.2387	0.852	0.1038	0.3357	0.758	0.3134	0.2118	0.139
$Q_L * LF$	-1.2529 ***	0.4792	0.009	-0.5198	0.7011	0.459	-0.9848 **	0.4560	0.031
$Q_L * SL$	0.0932	0.1812	0.607	-0.0084	0.2874	0.977	0.0952	0.1610	0.554
$Q_L * JET$	-0.2995	0.3119	0.337	-0.1378	0.5324	0.796	-0.0963	0.2632	0.714
$Q_L * SIZE$	0.1159	0.1909	0.544	0.1047	0.2912	0.720	0.0266	0.1714	0.877
$Q_K * LF$	1.2675 **	0.6688	0.058	1.1587	0.9225	0.211	0.7433	0.6395	0.245
$Q_K * SL$	0.2430	0.1930	0.208	0.6086 *	0.3117	0.053	0.0294	0.1773	0.868
$Q_K * JET$	0.0934	0.3227	0.772	0.2476	0.5345	0.644	0.0685	0.2710	0.801
$Q_K * SIZE$	-0.3233	0.2184	0.139	-0.4501	0.3476	0.197	-0.1541	0.1829	0.399
$LF^2/2$	-0.2952	1.9220	0.878	-0.7059	2.5419	0.782	0.0693	1.9108	0.971
LF*SL	0.1669	0.7224	0.817	0.7364	1.1108	0.508	-0.1252	0.6683	0.851
LF*JET	-1.3066	1.1015	0.236	-0.6086	1.6222	0.708	-2.1325 **	1.0320	0.039
LF*SIZE	2.1381 ***	0.7714	0.006	1.0604	1.0125	0.296	2.2467 ***	0.7231	0.002
$SL^2/2$	-0.3187	0.2814	0.257	-1.1233	0.6824	0.102	0.2830	0.2175	0.193
SL*JET	0.5056	0.6801	0.457	0.6476	1.4195	0.649	-0.2578	0.5127	0.615
SL*SIZE	-0.2258	0.3277	0.491	0.3829	0.4871	0.433	-0.4680	0.2912	0.108
$JET^2/2$	-0.8982	1.5261	0.556	-2.8892	2.7732	0.299	0.6862	1.3207	0.603
JET*SIZE	-0.1954	0.3535	0.580	-0.0562	0.5124	0.913	0.1144	0.3446	0.740
$SIZE^2/2$	-0.1791	0.4180	0.668	-1.0108	0.6392	0.116	0.2608	0.3503	0.457
σ_ε	0.1228 ***	0.0146	0.000	-	-	-	-	-	-
σ_e	0.1075 ***	0.0200	0.000	-	-	-	-	-	-
Constant	-	-	-	-1.2421	5.7035	0.828	-	-	-

Note: P-val<0.10: * , P-val<0.05: **, and P-val<0.01: ***

Figure 2: Efficiency Estimates for Translog Model
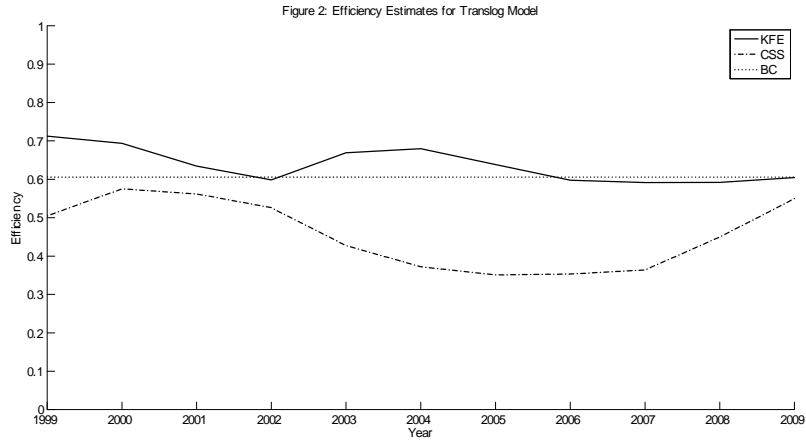
Now, we present the efficiency estimates when the trimming for KFE and CSSW are done for top-bottom 5% (rather than 7.5%) of the effects term when calculating the efficiencies. The BC estimates remain the same as they are not subject to such trimming. The average efficiency estimates for 5% trimming case are provided in Figure 3 and Figure 4.
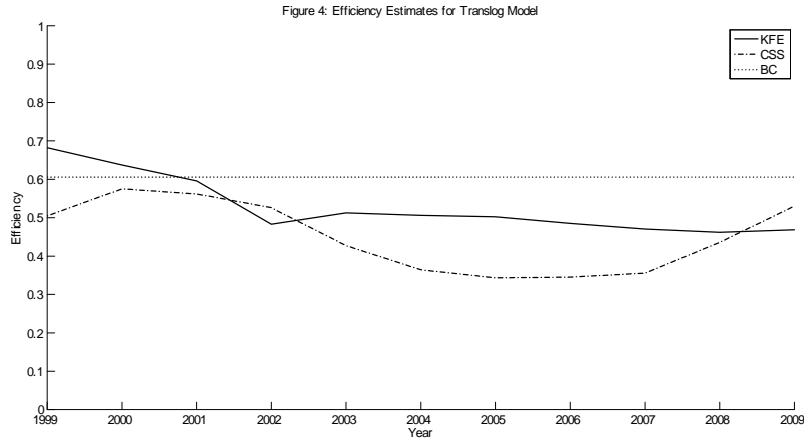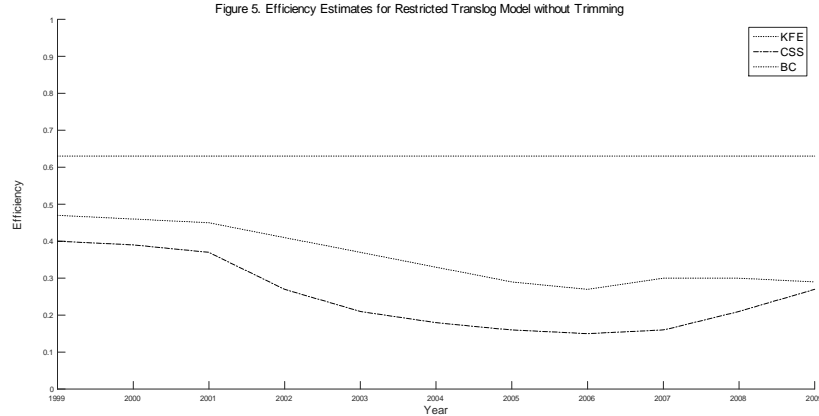

Figure 3: Efficiency Estimates for Restricted Translog Model

Figure 4: Efficiency Estimates for Translog Model
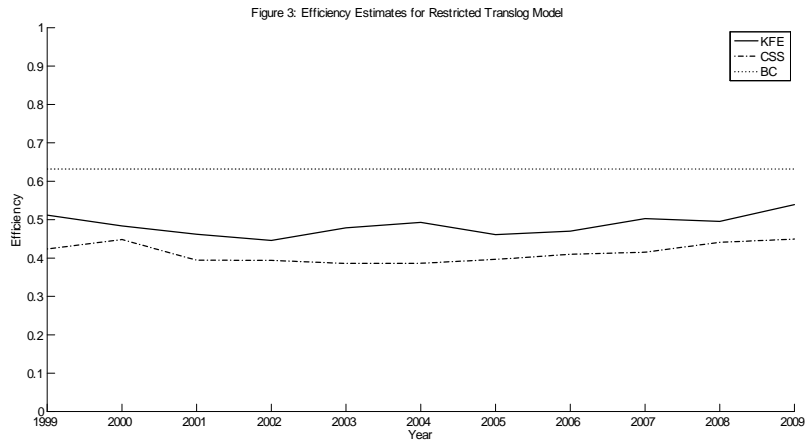
Figure 5. presents the average of efficiency estimates for restricted translog model. Due to outliers the KFE and CSS model estimates are low.

Figure 5. Efficiency Estimates for Restricted Translog Model without Trimming

Finally, we provide our estimates for alternative Kalman filter models in Table 11. As we mentioned we failed to estimate the full Kalman filter model that we presented in Equation 1. We rather estimated the models given in Equation 12 (random walk model with deterministic trend) and Equation 13 (AR(1) model).[21] Based on the BIC values

---

[21]The initial values for the AR(1) model are estimated as parameters.

Table 11. Alternative Kalman Filter Model Estimates

	Random Walk with Deterministic Trend				AR(1)			
	Param.		SE	P-val	Param.		SE	P-val
Q_L	0.5537		0.3573	0.1212	0.1723		0.2971	0.562
Q_K	1.7832	***	0.6854	0.0093	0.4194		0.3262	0.1986
LF	-5.8284	*	3.0391	0.0551	-6.5806	***	2.1545	0.0023
SL	3.0405		1.8894	0.1076	1.4127	**	0.6682	0.0345
JET	0.377		4.6739	0.9357	0.7883		2.4007	0.7426
SIZE	1.8449	*	0.9941	0.0635	1.1706		0.731	0.1093
$Q_L^2/2$	0.0664		0.1214	0.5843	0.0573		0.0719	0.426
$Q_L Q_K$	-0.2007		0.1751	0.2519	-0.0869		0.1047	0.4065
$Q_K^2/2$	0.0844		0.2344	0.7188	0.2174		0.1719	0.206
$LF^2/2$	1.3581		1.9264	0.4808	1.9357		1.7737	0.2751
LF*SL	0.0767		0.836	0.9269	-0.1756		0.6559	0.7889
LF*JET	-0.9579		1.3006	0.4614	-1.4287		0.9906	0.1492
LF*SIZE	1.2606		0.8475	0.1369	1.8966	***	0.7116	0.0077
$SL^2/2$	-0.3812		0.315	0.2262	0.0339		0.1509	0.8221
SL*JET	0.3743		0.8704	0.6672	0.3379		0.4714	0.4735
SL*SIZE	-0.1095		0.3146	0.7279	-0.2671	**	0.1315	0.0423
$JET^2 2$	-0.836		1.8501	0.6514	-0.5862		1.2679	0.6438
JET*SIZE	-0.2047		0.3968	0.606	-0.2582		0.2774	0.352
$SIZE^2 2$	-0.3137		0.4115	0.4459	0.0151		0.1779	0.9324
σ_ε	0.0925	***	0.0212	0	0.1703	***	0.0061	0
σ_e	0.1636	***	0.0242	0	0		0.0047	0.9997
ρ	-		-	-	1.0315	***	0.0254	0

Note: P-val<0.10: * , P-val<0.05: **, and P-val<0.01: ***

the random walk model with deterministic trend is not preferred. In particular, BIC values for random walk with deterministic trend and random walk models are 2.1 vs 1.7, respectively. The second model seems to be subject to the pile up problem as one of the variance parameters is collapsed to zero. Hence, we only provide the results for the sake of completeness. Nevertheless, the estimate of $\rho = 1.0315$ parameter indicates that our random walk assumption for the trend term is sensible for our empirical example. Hence, these findings support our choice for using the random walk model as our benchmark model.