Lancaster University Management School

Department of Economics

Essays on Stochastic Frontier Models

A Thesis Submitted for the Degree of

Doctor of Philosophy in Economics

Mariol Jonuzaj

June 16, 2023

Declaration

I hereby declare that this thesis is my own work and has not been submitted for the award of a higher degree elsewhere. This thesis contains no material previously published or written by any other person except where references have been made.

Mariol Jonuzaj
16/06/2023

Acknowledgments

First and foremost, I am greatly indebted to my supervisors, Professor Efthymios Tsionas and Dr Byunghoon Kang, for their support, guidance and advice. This thesis would not have been possible without their encouragement and insights. In addition, I would like to thank Professor Marwan Izzeldin for his support and the numerous insightful conversations we shared. Also, I would like to thank Dr. Ruijun Bu for acting as external examiner and for his valuable comments and suggestions greatly improved this dissertation.

I want to express my gratitude to the Lancaster University Management School and the Department of Economics for supporting me financially and giving me the opportunity to pursue this journey. Moreover, I want to thank Professor Efthymios Pavlidis, Dr Orestis Troumpounis, Dr Olivier Cardi, Dr David Rietzke, Dr Giorgio Motta and Dr Spyridon Lazarakis for their advice throughout these years. Special thanks go to Caren Wareing for her administrative support.

Also, I want to thank my PhD colleagues and friends Amanda De Pirro, Aristotelis Margaris, Chiara Costi and Wachid Muslimin for all the fun times we had together. I wish them all the best. Thank you to my girlfriend Elitsa Atanasova, for all her love and encouragement throughout this journey.

Last but not least, I want to thank my parents and my sister for their endless love and support.

Abstract

This thesis consists of three independent chapters. The first chapter "**Bayesian Inference in Dynamic Panel Stochastic Frontier Models**" proposes a new stochastic frontier model which accounts for the intertemporal production behaviour. The conceptualization is based on the notion that firms face production adjustment costs in the short run due to the presence of quasi-fixed inputs. Consequently, this sluggish adjustment of the entire production process will create a dependency between the current and past production state. To capture this dynamic process, this chapter utilizes the traditional partial adjustment mechanism. The mechanism delivers a dynamic specification and allows factor inputs and inefficiency shocks to have an intertemporal effect on the production process. Moreover, the model allows heterogeneous adjustment speeds and input elasticities across the production units. Model inference is based on Bayesian MCMC techniques with data-augmentation. We illustrate the new model in an empirical application where we estimate the productivity and efficiency growth of the Egyptian private manufacturing sector during the early 90's.

In a similar vein, the second chapter, "**Dynamic Panel Stochastic Frontier Models with Inefficiency Effects**", deals with dynamic panel frontier models where inefficiency effects can be a function of exogenous environmental variables. This chapter builds upon advancements in the field and utilizes parametric cumulative distribution functions to specify technical efficiency. The proposed model allows the presence of fixed effects and time-varying inefficiencies. Model estimation is based on the Generalized Method of Moments (GMM) approach, where various forms of input endogeneity can be effectively addressed.

Last, the third chapter "**A simple method for modelling the energy efficiency rebound effects with an application to energy demand frontiers**" proposes a new simple method for estimating the energy inefficiency rebound effects. Model estimation is based on a two-stage approach. In the first stage, we argue in estimating a reduced form stochastic frontier model with country-specific inefficiency heteroscedastic effects. In the second stage, the energy efficiency rebound effects can be obtained effectively using moment-matching methods such as the GMM approach. We apply the proposed model on aggregate energy frontiers where we estimate the energy efficiency and the corresponding rebound effects for a balanced panel of OECD economies. The empirical results suggest an overall upward trend of energy efficiency scores. The energy rebound effects range from 28% to 92%, indicating that energy efficiency actions could have a limited impact on achieving environmental objectives.

Contents

Chapter 1

Bayesian Inference in Dynamic Panel Stochastic Frontier Models

Abstract

In this chapter, we propose a dynamic panel stochastic frontier model that incorporates firms' intertemporal decision behaviour and the short-run stagnant adjustment of the production process. The dynamic production specification utilizes the fact that in the short-run, production systems can be subject to adjustment costs, and the final produced output can only be partially adjusted to the optimum production level. The proposed model nests previous panel stochastic frontier models and is able to adequately separate the firm-specific unobserved effects from the latent time-varying technical inefficiencies. Model inference is based on Bayesian Markov Chain Monte Carlo (MCMC) techniques with data augmentation. Using artificial data, we illustrate that our model performs very well in small and moderate samples. Last, we display the new model in an empirical example and illustrate that static specifications that ignore production adjustment effects can generate biased technical efficiency estimates. Overall, the proposed reduced-form model provides meaningful economic results that align with the recent macroeconomic and microeconomic literature, indicating that the proposed model can be considered a good alternative for economic analysis.

1.1 Introduction

Stochastic frontier analysis (SFA) was first introduced by Aigner et al. (1977) and Meeusen and van Den Broeck (1977) and incorporates the fact that the final production output can deviate from the maximum feasible output (production frontier) not only due to technical inefficiencies that may exist in the process but also due to stochastic factors outside the firm's control (e.g. general macroeconomic environment). Since then, SFA consists as a standard econometric tool for measuring production technical efficiencies and has been applied in various fields such as the neoclassical production theory, energy economics, banking economics, growth models, educational and health economics.

Despite the vast expansion of the stochastic frontier literature[1] that have emerged throughout the years, the main body of stochastic frontier models is built under different static frameworks. The primary assumption under a static specification is that firms can adjust their input levels costless, and their operational process can be instantaneously adjusted to the new economic conditions. Nevertheless, in real production environments, it is well documented that firms face factor adjustment costs, and hence, the intertemporal production decision is subject to those input adjustment constraints. In practical terms, adjustment costs can be present due to the fact that firms face short-run costs related to quasi-fixed costs. Some early influential papers that discuss the underlying economic mechanisms can be found in Lucas Jr (1967), Treadway (1970), Hamermesh (1995), Hamermesh and Pfann (1996), Nickell (1996). In this line, some more recent studies that illustrate the importance of accounting for the input adjustment frictions in the economic modelling consist of Hall (2004) where the author examines the factor adjustment costs for the US sectors, Cooper and Haltiwanger (2006) where capital adjustment costs are investigated in a sample of plant-level data, Groth and Khan (2010) examine the impact of investment costs on the US economy, Bergeaud and Ray (2021) explore the effects of real estate friction on firms' dynamics, and Artuç et al. (2010), Artuc et al. (2022) illustrate the effects of trade shocks on labour adjustment frictions, among others.

Said that it is apparent that in any production process, even in the absence of any other source of output fall, such as the presence of technical inefficiencies, the sluggish behaviour of factor inputs can directly affect the production process and, consequently, the final produced output. As a result, model and econometric specifications which ignore the aforementioned production dynamics can generate misleading inferences regarding the economic performance of a production unit.

[1] For an excellent literature review, see the textbooks by Kumbhakar and Lovell (2000), Greene (2008) and Kumbhakar et al. (2020) for more recent advances.

In this line, this paper introduces a reduced-form econometric model which is able to account for the evidence introduced above. Specifically, we present a generalized panel stochastic frontier model incorporating the short-run production dynamics arising from firms facing factor adjustment costs. We utilize a partial adjustment mechanism where the production output follows a sluggish intertemporal path towards the long-run equilibrium state. The proposed model specification delivers a dynamic panel model with time-varying inefficiency effects. In order to estimate the model and retrieve the latent technical inefficiency effects, we propose a Bayesian hierarchical model where the heterogeneity effects are incorporated using a random coefficient setting. The proposed specification allows for a flexible form of heterogeneity where firms can display different production capabilities. In addition, our approach allows economic sectors to face different speeds of adjustments and hence different intertemporal paths towards the equilibrium state. This specification allows us to analyze economic sectors more realistically since it is evident they are subject to different underlying economic conditions; therefore, each sector can replace its input factors at a different cost and pace.

Model inference and estimation are based on Bayesian Markov Chain Monte Carlo (MCMC) techniques with data augmentation. In particular, we demonstrate a Gibbs Sampler iteration approach where the corresponding marginal posterior distributions can be effectively obtained. Using generated artificial data, we test the performance of the proposed Gibbs Sampler and we illustrate that our Bayesian technique performs very well in both small and moderate samples.

Last, we apply our proposed model to the same dataset as in Getachew and Sickles (2007) and Bhattacharyya (2012), where the authors analyze the effects of market liberation during the early 90s in the Egyptian manufacturing sector. We illustrate that our model can adequately control for the unobserved heterogeneity effects and separate the manufacturing specif effects from the technical efficiency scores. Moreover, we show that when ignoring the presence of adjustment costs, the model will underestimate the technical efficiency scores. In addition, we illustrate the adjustment rate towards the optimal state is approximately 86%, indicating that labour mobility had a mild impact on the manufacturing sector. Overall, the proposed reduced-form model provides meaningful economic results that align with the recent macroeconomic and microeconomic literature, indicating that the proposed model can be considered a good alternative for economic analysis.

The rest of the chapter is organized as follows. In section 1.2, we present the theoretical framework of the partial adjustment effect and the proposed model. In section 1.3, we illustrate a generalization of the model. In section 1.4, we illustrate the proposed Bayesian approach. In section 1.5, we test the performance of the model using artificial data. In section 1.6, we present the empirical application. Last, in section 1.7, we conclude.

Related Literature: The proposed model fits into several different literature strands as we extend existing specifications and bring several distinct threads into the literature together. First, Nickell (1996) and Nickell et al. (1997) are some of the existing papers in the literature that utilize dynamic panel production functions to incorporate the short-run adjustment frictions and examine the firms' total factor productivity. In all these studies, the authors analyze the total factor productivity through control variables that affect productivity evolution and do not explicitly estimate the corresponding technical inefficiencies. Moreover, Ayed-Mouelhi and Goaied (2003) and Bhattacharyya (2012) utilize the partial adjustment mechanism to model the intertemporal production behaviour allowing the presence of time-invariant technical inefficiency. As in our model, they utilize a dynamic panel model where the autoregressive parameter of the lag output captures the magnitude of the sluggish adjustment. They propose a two-stage solution, wherein the first step, a Generalized Method of Moments (GMM) approach, is used to estimate the parameters of interest. In order to retrieve the corresponding efficiency scores, the authors use a Schmidt and Sickles (1984) approach where the technical inefficiency is obtained by relative comparison with the "fully efficient" firm. In our model, in contrast, we allow for a flexible specification where different sectors face different technological possibilities. In addition, we allow technical inefficiency to be time-varying, and we separate the unobserved heterogeneity effects from the inefficiency effects. This is quite important in practice since, as has been illustrated in the stochastic frontier literature, failing to control for the unobserved heterogeneity adequately will result in very distorted efficiency scores (see Greene (2005a,b), Wang and Ho (2010), Chen et al. (2014), Belotti and Ilardi (2018), Kutlu et al. (2019), among others.). This is evident in our empirical findings, where our model's absolute technical efficiency estimates are substantially higher than the relative efficiency scores.

Another strand of the literature deals with reduced-form dynamic models, where the technical inefficiency is specified as an autoregressive function of its past values. The motivation behind the autoregressive structure of technical inefficiency is that the input adjustment costs will cause sluggish adoption of new technological innovations. Therefore, the technical inefficiency evolution towards the long-run state will be more stagnant. Econometric modeling in this direction can be seen at Ahn et al. (2000), Tsionas (2006), Emvalomatis (2012a), Galán et al. (2015), Amsler et al. (2014) and Lai and Kumbhakar (2020).

Last, Tsionas et al. (2020) provide a new structural economic model where the production process can be subject to adjustment costs. Their framework allows technical inefficiencies to be determined endogenously from the firm's intertemporal optimization problem. Although structural models can provide comprehensive economic results, in this paper, we focus on providing a flexible reduced-form model that empirical scholars can easily implement.

1.2 Model Setup

To illustrate the idea of dynamic production frontiers, let's assume an economic sector where N firms are operating over a time period T. Each firm is using a set of inputs $\boldsymbol{X}_{it} \in \mathbb{R}_{+}^{K}$ in order to produce a single output $Y_{it}^{*} \in \mathbb{R}_{+}$. Output Y_{it}^{*} is a latent variable and it can be seen as the targeted level of output that the firm wants to achieve. This production process can be simply described as:

$$Y_{it}^{*} = f(\boldsymbol{X}_{it}; \boldsymbol{\beta}^{*})\boldsymbol{\tau}_{it}^{*} \tag{1.1}$$

where $f(.)$ can be any production function, such as the Cobb-Douglas or the Translog function, $\boldsymbol{X}_{it}$ is the set of inputs (e.g. labour, capital, energy, materials) used by firm i at time t, $\boldsymbol{\beta}^{*}$ is the vector of the technological parameters and $\boldsymbol{\tau}_{it}^{*}$ depicts the output-oriented technical efficiency of firm i at time t. Technical efficiency lies on the interval $\boldsymbol{\tau}_{it}^{*} \in (0, 1]$ and reflects the fact that firms may not lie on the production frontier, but instead, they can display technical inefficiencies which will deviate their production output from the frontier. The production model in 1.1 can be seen as the long-run equilibrium specification. Here, we should highlight that the specification in 1.1 allows firms to be technical inefficient even in the long run. The rationale behind this is that the firms can still survive and continue operating in a market even if they exhibit some level of technical inefficiency.

In logarithm form, the stochastic production frontier will be:

$$y_{it}^{*} = a_i^{*} + \boldsymbol{x}_{it}'\boldsymbol{\beta}^{*} + v_{it}^{*} - u_{it}^{*} \tag{1.2}$$

where y_{it}^{*} is the desired production log-output of firm i at time t, a_i^{*} is the usual firm-specific term which captures time-invariant heterogeneity across the firms, $\boldsymbol{x}_{it}$ is the $K \times 1$ vector of log-inputs used in the process, $\boldsymbol{\beta}^{*}$ is the $K \times 1$ parameter vector which can be interpreted as input elasticities, v_{it}^{*} is the symmetric error term which captures common statistical measurement errors and u_{it}^{*} is a non-negative term which determines the level of technical inefficiency of firm i at time t.

As mentioned above, firms due to different market conditions or the presence of input adjustment imperfections are not able to adjust their production instantaneously to their targeted production y_{it}^{*}. Instead, the final produced output faces a gradual adjustment towards the desired level. The simplest partial adjustment mechanism can be described as:

$$y_{it} - y_{it-1} = \lambda(y_{it}^{*} - y_{it-1}) \quad , \quad i = 1, 2, \ldots, N \quad , \quad t = 1, 2, \ldots, T \tag{1.3}$$

where y_{it} is the observed output at time t, y_{it-1} is the output produced at the lagged period $t-1$, y^*_{it} is the desired or targeted output at time t and $0 < \lambda \leq 1$ is the adjustment coefficient which reflects the speed of output adjustment towards the long-run level y^*_{it}. It is clear that in the case where $\lambda = 1$, the real output y_{it} is adjusted instantaneously to the desired level y^*_{it}. On the other hand, when $0 < \lambda < 1$, the process is subject to some level of inertia and the final output y_{it} lags behind the target level y^*_{it}. In the economics literature, specification in equation 1.3 is known as the Partial Adjustment Model and has been introduced by Nerlove (1958).

Combining the equilibrium relationship in equation 1.2 with the short-run dynamics from equation 1.3 , we obtain the final form of the model:

$$y_{it} = \rho y_{it-1} + \boldsymbol{x}'_{it}\boldsymbol{\beta} + a_i + v_{it} - u_{it} \tag{1.4}$$

where $\rho = (1-\lambda)$, $\boldsymbol{\beta} = \lambda\boldsymbol{\beta}^*$, $a_i = \lambda a^*_i$, $v_{it} = \lambda v^*_{it}$ and $u_{it} = \lambda u^*_{it}$. In this form, it is clear that the partial adjustment mechanism delivers a dynamic panel data stochastic frontier model. Estimation and inference in dynamic panel data models have been at the heart of modern econometrics and typically involve GMM methods. Relative literature consists of Holtz-Eakin et al. (1988), Anderson and Hsiao (1981, 1982), Arellano and Bond (1991) and Blundell and Bond (1998), where the authors proposed different GMM and SYS-GMM techniques. Recently, Cave et al. (2022) illustrated an extensive Monte Carlo simulation study, where the statistical performance of different dynamic panel estimators is evaluated.

Although the aforementioned GMM methods have been extensively used by empirical researchers, and the fact that estimation routines have been developed in most of the statistical software, in a model where the time-varying technical inefficiency term is incorporated, the above-stated literature cannot be useful. As stated above, by using the IV/GMM/SYS-GMM approach, we would be able to identify only relative inefficiency scores[2] and not the efficiency scores in absolute levels. Therefore, in this framework, it is obvious that introducing a dynamic panel stochastic frontier model which incorporates distribution assumptions for the time-varying inefficiency effects and at the same time controlling for the unobserved firm heterogeneity effects is of great importance.

1.3 A Generalization of the Model

To generalize the model in 1.4, one way to capture unobserved heterogeneity in the production process is to introduce a random coefficients stochastic frontier model where the heterogeneity

[2]Here we need to assume that at least one firm lies on the frontier (fully efficient firm) and the inefficiency scores of the remaining firms can be estimated relative to the best-performed firm.

effects are not only captured by the usual time-invariant random effects term but also from the fact that different firms are facing different technological capabilities. The fact that different firms face different production capabilities has been evident in various empirical examples where random coefficients settings have been employed (see Tsionas (2002), Emvalomatis (2012b), Feng et al. (2018), Tsionas and Tzeremes (2021), among others). Moreover, in a dynamic panel setting similar evidence has been illustrated by Hsiao et al. (1999), Liu et al. (2017), Assaf and Tsionas (2019). Our proposed model is based on the dynamic panel model with heterogeneous effect proposed by Hsiao et al. (1999). The authors illustrate that the Bayesian hierarchical approach is asymptotically equivalent to the mean group estimator of Pesaran and Smith (1995) as $N \to \infty$ and $T \to \infty$, and illustrate that the hierarchical model performs better in small and moderate samples.

Hence, the proposed generalized panel stochastic frontier model has the form:

$$y_{it} = \rho_i y_{it-1} + \boldsymbol{x}'_{it}\boldsymbol{\beta}_i + a_i + v_{it} - u_{it} \tag{1.5}$$

Throughout the paper, we make the following model assumptions:

Assumption 1 *The common measurement errors $v_{it} \sim iid(0, \sigma_v^2)$.*

Assumption 2 *The inefficiency term u_{it} is iid from a probability density distribution with non-negative support.*

Assumption 3 *The measurement errors v_{it} are independent of the inefficiency term u_{it}.*

Assumption 4 *The input vector $\boldsymbol{x}_{it}$ is independent of the error term v_{is} and the inefficiency term u_{is} for all t and s.*

Assumption 1 is common in the stochastic frontier literature. **Assumption 2** derives naturally from the fact that the inefficiency term cannot take negative values. In addition, **Assumption 3** is dominant in the stochastic frontier literature and is vital for the derivation of the marginal distribution of the composite term $\varepsilon_{it} = v_{it} - u_{it}$. Last, **Assumption 4** implies that the input vector is strictly exogenous. In general, we think that all the above assumptions can be considered minimalistic in order to identify the parameters of interest as long as the latent inefficiency scores.

In matrix form, the above model can be written as:

$$\boldsymbol{y}_i = \rho_i \boldsymbol{y}_{i,-1} + \boldsymbol{X}_i\boldsymbol{\beta}_i + \boldsymbol{v}_i - \boldsymbol{u}_i \text{ , for each } i = 1, 2, \ldots, N \tag{1.6}$$

where $\boldsymbol{y}_i = [y_{i1}, y_{i2}, \ldots, y_{iT}]'$ is a $T \times 1$ matrix of the dependent variable, $\boldsymbol{y}_{i,-1} = [y_{i0}, y_{i1}, \ldots, y_{iT-1}]'$ is a $T \times 1$ matrix of the lagged log-output, $\boldsymbol{X}_i = [\boldsymbol{x}'_{i1}, \boldsymbol{x}'_{i2}, \ldots, \boldsymbol{x}'_{iT}]'$ is a $T \times K$ matrix of the log-inputs, $\boldsymbol{v}_i = [v_{i1}, v_{i2}, \dots, v_{iT}]$ is the $T \times 1$ matrix of the symmetric error term and $\boldsymbol{u}_i = [u_{i1}, u_{i2}, \dots, u_{iT}]'$ is the $T \times 1$ matrix of the technical inefficiency term. Regarding the initial values conditions y_{i0} for each $i = 1, 2, \dots, N$, we depart from the approach used in Hsiao et al. (1999), where the initial values are considered as fixed and known, and instead, we are treating them as latent variables to be estimated.

In order to complete the model, for the two-sided error term we assume $\boldsymbol{v}_i \sim \boldsymbol{\mathcal{N}}(0, \sigma_v^2)$ and for the inefficiency term, we make use of the half-normal distribution, as proposed by Aigner et al. (1977)[3], with:

$$\boldsymbol{u}_i \sim \boldsymbol{\mathcal{N}}^+(0, \sigma_u^2) \tag{1.7}$$

It is obvious that the proposed model in 1.5, can be seen as a generalization of previous panel data stochastic frontier models that have been introduced in the literature, so far. First, we should note that in the case where firms adjust perfectly their production to their targeted levels, viz. $\rho_i = 0$, we obtain the stochastic frontier model with random coefficients as introduced by Tsionas (2002). As discussed in the original paper, the models account for the fact that different firms may face different technological capabilities. In addition, when the random coefficients assumption is constrained to a fixed coefficient specification, viz $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \cdots = \boldsymbol{\beta}_N = \bar{\boldsymbol{\beta}}$, then the random coefficient SFM returns to a stochastic frontier model where the only source of heterogeneity is captured by the usual firm specific a_i term. This model specification is known as the True Fixed Effects (TFE) or True Random Effects (TRE) model and was originally proposed by Greene (2005a,b)[4]. Both specifications are able to separate the firm-specific heterogeneity effects from the firm's technical inefficiency. In addition, Fernandez et al. (1997) present different panel stochastic frontier models using Bayesian posterior inference.

Last, in the case of a simple stochastic frontier model under the absence of any heterogeneity source, viz $\sigma_a^2 = 0$, we arrive at the original model as proposed by Aigner et al. (1977). Under the Bayesian framework, estimation inference for the normal-half-normal model was presented by Van den Broeck et al. (1994) and later by Tsionas (2001)[5].

[3]Other well-known specifications, include the Exponential function proposed by Meeusen and van Den Broeck (1977), the Truncated Normal distribution proposed by Stevenson (1980) and the Gamma distribution introduced by Greene (1990).

[4]A parallel literature which elaborates panel stochastic frontier models under fixed effects consists of Wang and Ho (2010), Chen et al. (2014), Belotti and Ilardi (2018), Kutlu et al. (2019), among others.

[5]In both papers, the authors present posterior inference under a truncated normal distribution with $u \sim N^+(\mu, \sigma_u^2)$. Obviously, the Normal-Half-Normal model can obtained by imposing $\mu = 0$.

1.4 Bayesian Estimation

In this section, we built a Bayesian Hierarchical model in order to estimate consistently the parameters of interest. For illustration purposes, we rewrite the model in 1.6, as:

$$\boldsymbol{y}_i = \boldsymbol{Z}_i \boldsymbol{\delta}_i + \boldsymbol{v}_i - \boldsymbol{u}_i \text{ , for each } i = 1, 2, \ldots, N \tag{1.8}$$

with

$$\boldsymbol{v}_i \sim \boldsymbol{\mathcal{N}}(0, \sigma_v^2) \quad , \quad \boldsymbol{u}_i \sim \boldsymbol{\mathcal{N}}^+(0, \sigma_u^2)$$

where $\boldsymbol{Z}_i = [\boldsymbol{y}_{i,-1}, \boldsymbol{X}_i]$ will be a $T \times J$ matrix and $\boldsymbol{\delta}_i = [\rho_i, \boldsymbol{\beta}_i]'$ is the $J \times 1$ matrix of unknown parameters, where $J = K + 1$. Equation 1.8 consists of the first stage of our hierarchical structure.

1.4.1 Hierarchical Priors

Having defined the first stage of the model, we need to proceed with the following stages of our hierarchical model. For the heterogeneous frontier coefficients, a convenient hierarchical prior distribution is to assume that each $\boldsymbol{\delta}_i$ is an independent draw from a multivariate Normal distribution, with:

$$\boldsymbol{\delta_i} = \begin{bmatrix} \rho_i \\ \boldsymbol{\beta_i} \end{bmatrix} \sim \boldsymbol{\mathcal{N}}(\bar{\boldsymbol{\delta}}, \boldsymbol{\Omega}) \tag{1.9}$$

where $\bar{\boldsymbol{\delta}}$ is a $J \times 1$ matrix of the mean values of the parameters of interest, and $\boldsymbol{\Omega}$ is a general positive-definite $J \times J$ variance-covariance matrix of the parameters. In addition, instead of a general covariance matrix, one can restrict the assumption of correlated parameters and assume that $\boldsymbol{\Omega}$ is a diagonal matrix, which implies that the parameters are independent of each other. The third stage of the hierarchical structure for the frontier parameters is to assume again a multivariate Normal distribution of the form:

$$\bar{\boldsymbol{\delta}} \sim \boldsymbol{\mathcal{N}}(\boldsymbol{\mu_0}, \boldsymbol{\Lambda_0}) \tag{1.10}$$

where $\boldsymbol{\mu_0}$ is a $J \times 1$ vector of mean values and $\boldsymbol{\Lambda_0}$ is the corresponding $J \times J$ variance-covariance matrix, which both of them are assigned by the researcher, according to her prior beliefs. From 1.9, it is obvious that one can assume a fully non-informative flat prior by imposing $\boldsymbol{\mu_0} = \boldsymbol{0}_{J \times 1}$ and $\boldsymbol{\Lambda_0} = 10^3 \times \boldsymbol{I}_{J \times J}$, where $\boldsymbol{0}_{J \times 1}$ is the $(J \times 1)$ null matrix and $\boldsymbol{I}_{J \times J}$ is a $(J \times J)$ identity matrix.

For the prior distribution of the variance-covariance matrix $\boldsymbol{\Omega}$, we assume an Inverse-Wishart conjugate prior distribution of the form:

$$\boldsymbol{\Omega} | \boldsymbol{\Psi}_0, v_0 \sim \boldsymbol{\mathcal{IW}}(\boldsymbol{\Psi}_0, v_0) \tag{1.11}$$

with probability density function:

$$\pi(\boldsymbol{\Omega}|\boldsymbol{\Psi}_0, v_0) \propto det(\boldsymbol{\Omega})^{-\frac{v_0+J+1}{2}} exp\left\{-\frac{1}{2}tr(\boldsymbol{\Psi}_0\boldsymbol{\Omega}^{-1})\right\} \tag{1.12}$$

where $det(.)$ corresponds to the determinant of a matrix, $tr(.)$ denotes the trace of a matrix, $\boldsymbol{\Psi}_0$ is the $J \times J$ positive definite scale matrix and v_0 are the corresponding degrees of freedom. Obviously, the parameter v_0 and the scale matrix $\boldsymbol{\Psi}_0$ are assigned by the researcher reflecting her prior belief on the variance-covariance matrix $\boldsymbol{\Omega}$. Here, a common choice is to set $v_0 = 0$ and $\boldsymbol{\Psi}_0 = 10^{-6} \times \boldsymbol{I}_{J\times J}$, which leads to the multivariate non-informative Jeffrey's prior, of the form:

$$\boldsymbol{\Omega} \propto det(\boldsymbol{\Omega})^{-\frac{J+1}{2}} \tag{1.13}$$

In addition, since the mean of the distribution is given by:

$$E(\boldsymbol{\Omega}|\boldsymbol{\Psi}_0, v_0) = \frac{\boldsymbol{\Psi}_0}{v_0 - J - 1} \quad , \quad v_0 > J + 1$$

one can pick $v_0 = J + 2$ in order for the prior mean of the variance-covariance matrix to be free from the corresponding degrees of freedom. As a result, the prior elicitation for the expectation of $\boldsymbol{\Omega}$ will depend only on the choice of $\boldsymbol{\Psi}_0$. For instance, if the researcher believes that the variance of the parameters $\boldsymbol{\delta}_i$ is 0.1, then a natural prior would be $\boldsymbol{\Psi}_0 = 0.1 \times \boldsymbol{I_{J\times J}}$.

Turning now our attention to the case where the variance-covariance matrix $\boldsymbol{\Omega}$ is diagonal, with:

$$\boldsymbol{\Omega} = \begin{bmatrix} \omega_1^2 & 0 & 0 & \dots & 0 \\ 0 & \omega_2^2 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \omega_J^2 \end{bmatrix} \tag{1.14}$$

the choice of the Inverse-Wishart distribution can not still be valid, since it assumes some correlation between the parameters $\bar{\boldsymbol{\delta}}$. Instead, by incorporating a diagonal matrix with the probability density in equation 1.12, we obtain:

$$\pi(\boldsymbol{\Omega}|\boldsymbol{\Psi}_0, v_0) \propto \left(\prod_{j=1}^{J} \omega_j^2\right)^{-\frac{v_0+J+1}{2}} \prod_{j=1}^{J} exp\left\{-\frac{\frac{1}{2}\psi_{0j}}{\omega_j^2}\right\} \tag{1.15}$$

where ψ_{0j} is the j^{th} element of $\boldsymbol{\Psi}_0$. From equation 1.15, we see that the obtained probability density implies Inverse Gamma type prior distributions for each element ω_j^2 of the form:

$$\pi(\omega_j^2|\psi_{0j}, v_0) \propto (\omega_j^2)^{-\frac{v_0+J-1}{2}-1} exp\left\{-\frac{\frac{1}{2}\psi_{0j}}{\omega_j^2}\right\} \quad , \quad j = 1, 2, \dots, J \tag{1.16}$$

or more simply:

$$\omega_j^2|\psi_{0j}, v_0 \propto \mathcal{IG}\left(\frac{v_0 + J - 1}{2}, \frac{1}{2}\psi_{0j}\right) \quad , \quad j = 1, 2, \dots, J \tag{1.17}$$

For the variance of the error term, we use the Inverse-Gamma prior distribution, of the form:

$$\sigma_v^2 | a_0, a_1 \sim \mathcal{IG}(a_0, a_1) \tag{1.18}$$

where a_0 and a_1 correspond to the shape and the scale parameter, respectively. The choice of the Inverse-Gamma distribution is a standard approach in Bayesian econometrics since it is a natural conjugate prior distribution for the variance of a normal distribution. In the special case where $a_0 = 0$ and $a_1 = 0$, we obtain the standard non-informative Jefrey's prior, viz $\pi(\sigma_v^2) \propto \frac{1}{\sigma_v^2}$. However, in order for the posterior distribution to be well defined, the shape and the scale hyperparameters can not be zero. Instead, in order to achieve the same result, one can assign values close to zero, such as $a_0 = a_1 = 10^{-6}$ (see Fernandez et al. (1997), page 186).

A similar approach can be followed for the variance of the inefficiency term, where again, an Inverse-Gamma distribution can be assigned as a prior belief, with:

$$\sigma_u^2 | \gamma_0, \gamma_1 \sim \mathcal{IG}(\gamma_0, \gamma_1) \tag{1.19}$$

where γ_0 and γ_1 correspond to the shape and the scale parameter, respectively. Again, the two hyperparameters are assigned by the researcher according to her prior belief regarding the technical inefficiency scores.

Last, a convenient way to state prior "ignorance" regarding the initial values y_{i0} for each $i = 1, 2, \ldots, N$, is to assign a vague normal distribution prior of the form:

$$y_{i0} \sim \mathcal{N}(\bar{y}_0, \sigma_0^2) \tag{1.20}$$

This states that the latent initial values conditions are distributed around a common mean $\bar{y}_0$ and a variance σ_0^2. The above probability density function can be written as:

$$y_{i0} = \bar{y}_0 + \xi_i \quad , \quad \xi_i \sim \mathcal{N}(0, \sigma_0^2) \quad , \quad i = 1, 2, \ldots, N \tag{1.21}$$

Interestingly, such "knowing little" prior specification is aligned with the initial condition assumptions introduced in Anderson and Hsiao (1981, 1982). As stated by the authors, the above specification depicts the firm initial endowments and these effects gradually vanish over time.

1.4.2 Posterior Analysis

Following the above Bayesian hierarchical structure and using the Bayes rule, we have:

$$p(\boldsymbol{\theta}, \boldsymbol{\delta}_i, \boldsymbol{u}, \boldsymbol{y}_0 | \boldsymbol{Y}, \boldsymbol{Z}) \propto f(\boldsymbol{Y}, \boldsymbol{Z} | \boldsymbol{\theta}, \boldsymbol{\delta}_i, \boldsymbol{u}, \boldsymbol{y}_0) f(\boldsymbol{\delta}_i, \boldsymbol{u}, \boldsymbol{y}_0 | \boldsymbol{\theta}) \pi(\boldsymbol{\theta}) \tag{1.22}$$

where $\boldsymbol{Y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_N]'$ and $\boldsymbol{Z} = [\boldsymbol{Z}_1, \boldsymbol{Z}_2, \ldots, \boldsymbol{Z}_N]'$, $p(\boldsymbol{\theta}, \boldsymbol{\delta}_i, \boldsymbol{u}, \boldsymbol{y}_0 | \boldsymbol{Y}, \boldsymbol{Z})$ denotes the augmented posterior distribution of the structural parameters $\boldsymbol{\theta} = [\bar{\boldsymbol{\delta}}, \sigma_v^2, \sigma_u^2, \boldsymbol{\Omega}]'$ and the latent parameters $\boldsymbol{\delta}_i, \boldsymbol{u}, \boldsymbol{y}_0$, $f(\boldsymbol{Y}, \boldsymbol{Z}|\boldsymbol{\theta}, \boldsymbol{\delta}_i, \boldsymbol{u}, \boldsymbol{y}_0)$ depicts the conditional likelihood function of the observed data from the first stage of the hierarchical structural, $f(\boldsymbol{\delta}_i, \boldsymbol{u}, \boldsymbol{y}_0|\boldsymbol{\theta})$ illustrates the distribution assigned in the second stage of the model and $\pi(\boldsymbol{\theta})$ depicts the prior distribution of the unknown structural parameters $\boldsymbol{\theta}$.

Since, $\boldsymbol{\delta}_i, \boldsymbol{u}, \boldsymbol{y}_0$ are elements not observed by the researcher, we can not provide posterior inference about the structural parameter vector $\boldsymbol{\theta}$ conditioning on the latent parameters. For this reason, the unobserved elements have to be integrated out of the posterior density function in 1.22. Therefore, in order to obtain the marginal posterior distribution of the structural parameters $\boldsymbol{\theta}$, we need to integrate the augmented posterior distribution with respect to the latent parameters $\boldsymbol{\delta}_i$, $\boldsymbol{u}$ and $\boldsymbol{y}_0$ as:

$$p(\boldsymbol{\theta}|\boldsymbol{Y}, \boldsymbol{Z}) \propto \int\int\int p(\boldsymbol{\theta}, \boldsymbol{\delta}_i, \boldsymbol{u}, \boldsymbol{y}_0|\boldsymbol{Y}, \boldsymbol{Z}) d\boldsymbol{\delta}_i d\boldsymbol{u} d\boldsymbol{y}_0 \tag{1.23}$$

The marginal distribution in 1.23 involves high dimensional integration and the provision of an exact solution seems not to be feasible. For this reason, in order to resolve the high complexity of the marginal distribution, we can treat the latent elements of the model $\boldsymbol{\delta}_i, \boldsymbol{u}, \boldsymbol{y}_0$ as unknown parameters to be estimated and posterior inference can be conducted using Bayesian MCMC techniques such as the Gibbs Sampler. This method is called data augmentation and has been initially proposed by Tanner and Wong (1987)[6].

[6] Bayesian MCMC with data-augmentation consists a standard tool in Bayesian stochastic frontier econometrics.

1.4.3 The Joint Augmented Posterior

Using the above Bayesian hierarchical structure, the joint augmented posterior density of the unknown parameter vector $\boldsymbol{\Theta} = (\bar{\boldsymbol{\delta}}, \{\boldsymbol{\delta}_i\}_{i=1,2,\dots,N}, \sigma_v^2, \sigma_u^2, \boldsymbol{\Omega}, \{y_{i0}\}_{i=1,2,\dots,N}, \boldsymbol{u})'$, will be given by:

$$
\begin{aligned}
p(\boldsymbol{\Theta}|\boldsymbol{y}, \boldsymbol{Z}, \boldsymbol{\Psi}_0, v_0, \boldsymbol{\mu}_0, \boldsymbol{\Lambda}_0) \propto \\
& (\sigma_v^2)^{-\frac{NT}{2}} exp\left\{-\frac{1}{2\sigma_v^2}\sum_{i=1}^{N}(\boldsymbol{y}_i - \boldsymbol{Z}_i\boldsymbol{\delta}_i + \boldsymbol{u}_i)'(\boldsymbol{y}_i - \boldsymbol{Z}_i\boldsymbol{\delta}_i + \boldsymbol{u}_i)\right\} \\
& \times (\sigma_u^2)^{-\frac{NT}{2}} exp\left\{-\frac{1}{2\sigma_u^2}\sum_{i=1}^{N}\boldsymbol{u}_i'\boldsymbol{u}_i\right\} \mathbf{1}(\boldsymbol{u}_i \geq 0) \\
& \times det(\boldsymbol{\Omega})^{-\frac{N}{2}} exp\left\{-\frac{1}{2}\sum_{i=1}^{N}(\boldsymbol{\delta}_i - \bar{\boldsymbol{\delta}})'\boldsymbol{\Omega}^{-1}(\boldsymbol{\delta}_i - \bar{\boldsymbol{\delta}})\right\} \\
& \times det(\boldsymbol{\Lambda}_0)^{-\frac{1}{2}} exp\left\{-\frac{1}{2}(\bar{\boldsymbol{\delta}} - \boldsymbol{\mu}_0)'\boldsymbol{\Lambda}_0^{-1}(\bar{\boldsymbol{\delta}} - \boldsymbol{\mu}_0)\right\} \\
& \times (\sigma_v^2)^{-a_0-1} exp\left\{-\frac{a_1}{\sigma_v^2}\right\} \\
& \times (\sigma_u^2)^{-\gamma_0-1} exp\left\{-\frac{\gamma_1}{\sigma_u^2}\right\} \\
& \times det(\boldsymbol{\Omega})^{-\frac{J+v_0+1}{2}} exp\left\{-\frac{1}{2}tr(\boldsymbol{\Psi}_0\boldsymbol{\Omega}^{-1})\right\} \\
& \times (\sigma_0^2)^{-\frac{N}{2}} exp\left\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{N}(y_{i0} - \bar{y}_0)^2\right\}
\end{aligned} \tag{1.24}
$$

The first line of the joint posterior distribution corresponds to the likelihood function of $\boldsymbol{y}_i$ and arrives from the fact that $\boldsymbol{y}_i|\boldsymbol{Z}_i, \boldsymbol{u}_i \sim \boldsymbol{\mathcal{N}}(\boldsymbol{Z}_i\boldsymbol{\delta}_i - \boldsymbol{u}_i, \sigma_v^2)$ for each $i = 1, 2, \dots, N$. The second line determines the half-normal distribution of the technical inefficiency term $\boldsymbol{u}_i$. In addition to, the third equation reflects the normal multivariate structure of the heterogeneous coefficients $\boldsymbol{\delta}_i$. The remaining five equations depict the prior information assigned by the researcher for the parameters $\bar{\boldsymbol{\delta}}$, σ_v^2, σ_u^2, $\boldsymbol{\Omega}$ and $\boldsymbol{y}_0$, respectively.

1.4.4 Posterior Analysis using Gibbs Sampling

Gibbs sampling is an iterative procedure that utilizes the conditional density distributions of the unknown parameter vector $\boldsymbol{\Theta}$ from the joint posterior distribution in 1.24. The iterated algorithm is used to approximate the marginal distribution of the corresponding parameters. In order to obtain a sample $\{\boldsymbol{\Theta}^s$ with $s = 1, 2, \dots, S\}$ that converges to the marginal distribution of each unknown parameter, we can follow the below procedure:

- Step 1: Draw $\boldsymbol{\delta}_i \sim p(\boldsymbol{\delta}_i|\bar{\boldsymbol{\delta}}, \sigma_v^2, \sigma_u^2, \boldsymbol{\Omega}, \boldsymbol{u}, \boldsymbol{y}_0)$
- Step 2: Draw $\bar{\boldsymbol{\delta}} \sim p(\bar{\boldsymbol{\delta}}|\boldsymbol{\delta}_i\sigma_v^2, \sigma_u^2, \boldsymbol{\Omega}, \boldsymbol{u}, \boldsymbol{y}_0)$

- Step 3: Draw $\sigma_v^2 \sim p(\sigma_v^2|\bar{\boldsymbol{\delta}}, \boldsymbol{\delta}_i, \sigma_u^2, \boldsymbol{\Omega}, \boldsymbol{u}, \boldsymbol{y}_0)$
- Step 4: Draw $\sigma_u^2 \sim p(\sigma_u^2|\bar{\boldsymbol{\delta}}, \boldsymbol{\delta}_i, \sigma_v^2, \boldsymbol{\Omega}, \boldsymbol{u}, \boldsymbol{y}_0)$
- Step 5: Draw $\boldsymbol{\Omega} \sim p(\boldsymbol{\Omega}|\bar{\boldsymbol{\delta}}, \boldsymbol{\delta}_i, \sigma_v^2, \sigma_u^2, \boldsymbol{u}, \boldsymbol{y}_0)$
- Step 6: Draw $\boldsymbol{u}_i \sim p(\boldsymbol{u}_i|\bar{\boldsymbol{\delta}}, \boldsymbol{\delta}_i, \sigma_v^2, \sigma_u^2, \boldsymbol{\Omega}, \boldsymbol{y}_0)$
- Step 7: Draw $\boldsymbol{y}_0 \sim p(\boldsymbol{y}_0|\bar{\boldsymbol{\delta}}, \boldsymbol{\delta}_i, \sigma_v^2, \sigma_u^2, \boldsymbol{\Omega}, \boldsymbol{u})$
- Step 8: Repeat the above steps S times

Once we obtain the S sample from the conditional distribution of the parameters $\boldsymbol{\Theta}$, the posterior mean estimates can be obtained by:

$$\hat{\boldsymbol{\Theta}} = \frac{1}{S}\sum_{s=1}^{S} \boldsymbol{\Theta}^s$$

1.4.5 Conditional Posterior Distributions

In order to perform the Gibbs Sampler, knowledge of the conditional posterior distributions is required. The conditional posterior distributions of the parameters of interest are presented below:

- **Conditional Distribution of $\boldsymbol{\delta}_i$'s:**

$$\boldsymbol{\delta}_i|\bar{\boldsymbol{\delta}}, \sigma_v^2, \sigma_u^2, \boldsymbol{\Omega}, \mathbf{u}, \boldsymbol{y}_0 \sim \boldsymbol{\mathcal{N}}(\hat{\boldsymbol{b}}_i, \hat{\boldsymbol{V}}_i) \tag{1.25}$$

where

$$\hat{\boldsymbol{b}}_i = \left(\frac{\boldsymbol{Z}_i'\boldsymbol{Z}_i}{\sigma_v^2} + \boldsymbol{\Omega}^{-1}\right)^{-1}\left(\frac{\boldsymbol{Z}_i'(\boldsymbol{y}_i + \boldsymbol{u}_i)}{\sigma_v^2} + \boldsymbol{\Omega}^{-1}\bar{\boldsymbol{\delta}}\right)$$
$$\hat{\boldsymbol{V}}_i = \left(\frac{\boldsymbol{Z}_i'\boldsymbol{Z}_i}{\sigma_v^2} + \boldsymbol{\Omega}^{-1}\right)^{-1}$$

- **Conditional Distribution of $\bar{\boldsymbol{\delta}}$:**

$$\bar{\boldsymbol{\delta}}|\boldsymbol{\delta}_i\sigma_v^2, \sigma_u^2, \boldsymbol{\Omega}, \mathbf{u}, \boldsymbol{y}_0 \sim \boldsymbol{\mathcal{N}}(\boldsymbol{B}[N\boldsymbol{\Omega}^{-1}\tilde{\boldsymbol{\delta}} + \boldsymbol{\Lambda}_0^{-1}\boldsymbol{\mu}_0], \boldsymbol{B}) \tag{1.26}$$

where

$$\tilde{\boldsymbol{\delta}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\delta}_i$$
$$\boldsymbol{B} = [N\boldsymbol{\Omega}^{-1} + \boldsymbol{\Lambda}_0^{-1}]^{-1}$$

- **Conditional Distribution of σ_v^2:**

$$\sigma_v^2|\bar{\boldsymbol{\delta}}, \boldsymbol{\delta}_i, \sigma_u^2, \boldsymbol{\Omega}, \mathbf{u}, \boldsymbol{y}_0 \sim \mathcal{IG}\left(\frac{NT}{2} + a_0, \frac{\sum_{i=1}^{N}(\boldsymbol{y}_i - \boldsymbol{Z}_i\boldsymbol{\delta}_i + \boldsymbol{u}_i)'(\boldsymbol{y}_i - \boldsymbol{Z}_i\boldsymbol{\delta}_i + \boldsymbol{u}_i)}{2} + a_1\right) \tag{1.27}$$

- **Conditional Distribution of σ_u^2:**

$$\sigma_u^2|\bar{\boldsymbol{\delta}}, \boldsymbol{\delta}_i, \sigma_v^2, \boldsymbol{\Omega}, \mathbf{u}, \boldsymbol{y}_0 \sim \mathcal{IG}\left(\frac{NT}{2} + \gamma_0, \frac{\sum_{i=1}^N \boldsymbol{u}_i'\boldsymbol{u}_i}{2} + \gamma_1\right) \tag{1.28}$$

- **Conditional Distribution of $\boldsymbol{\Omega}$ (non-diagonal):**

$$\boldsymbol{\Omega}|\bar{\boldsymbol{\delta}}, \boldsymbol{\delta}_i, \sigma_v^2, \sigma_u^2, \mathbf{u}, \boldsymbol{y}_0 \sim \boldsymbol{\mathcal{IW}}\left(N + v_0, \sum_{i=1}^N (\boldsymbol{\delta}_i - \bar{\boldsymbol{\delta}})(\boldsymbol{\delta}_i - \bar{\boldsymbol{\delta}})' + \boldsymbol{\Psi}_0\right) \tag{1.29}$$

- **Conditional Distribution of the diagonal elements ω_j^2 of $\boldsymbol{\Omega}$ (in the case of diagonal matrix):**

$$\omega_j^2|\bar{\boldsymbol{\delta}}, \boldsymbol{\delta}_i, \sigma_v^2, \sigma_u^2, \mathbf{u}, \boldsymbol{y}_0 \sim \mathcal{IG}\left(\frac{N + v_0 + J - 1}{2}, \frac{\sum_{i=1}^N (\delta_{ij} - \bar{\delta}_j)^2 + \psi_{0j}}{2}\right) \tag{1.30}$$

- **Conditional Distribution of $\boldsymbol{u}_i$'s:**

$$\boldsymbol{u}_i|\bar{\boldsymbol{\delta}}, \boldsymbol{\delta}_i, \sigma_v^2, \sigma_u^2, \boldsymbol{\Omega}, \boldsymbol{y}_0 \sim \boldsymbol{\mathcal{N}}^+\left(-\frac{\sigma_u^2(\boldsymbol{y}_i - \boldsymbol{Z}_i\boldsymbol{\delta}_i)}{\sigma_u^2 + \sigma_v^2}, \frac{\sigma_v^2\sigma_u^2}{\sigma_u^2 + \sigma_v^2}\right) \tag{1.31}$$

- **Conditional Distribution of $\boldsymbol{y}_0$:**

The two equations involving the unobserved initial values y_{i0} are:

$$y_{i0} = \bar{y}_0 + \xi_i \quad , \quad \xi_i \sim N(0, \sigma_0^2) \quad , \quad i = 1, 2, \ldots, N$$

which derives from the prior normal probability density in equations 1.20-1.21, and

$$y_{i1} = \rho_i y_{i0} + x_{i1}'\beta_i + v_{i1} - u_{i1} \quad , \quad i = 1, 2, \ldots, N$$

which comes from the dynamic panel production frontier model at time $t = 1$. In matrix form, the above two equations can be written as:

$$\begin{bmatrix} y_{i1} - x_{i1}'\beta_i + u_{i1} \\ \bar{y}_0 \end{bmatrix} = \begin{bmatrix} \rho_i \\ 1 \end{bmatrix} y_{i0} + \begin{bmatrix} v_{i1} \\ \xi_i \end{bmatrix} \quad , \quad i = 1, 2, \ldots, N$$

where

$$\begin{bmatrix} v_{i1} \\ \xi_i \end{bmatrix} \sim \boldsymbol{\mathcal{N}}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_0^2 \end{bmatrix}\right)$$

Therefore, in the above form, we can treat the initial conditions y_{i0}, as the unknown parameter of a linear model. Using the Theil-Goldberger Generalized Linear Mixed estimator (see Theil and Goldberger (1961)), the initial values $\{y_{i0}\}_{i=1,2,\ldots,N}$ can be obtained from:

$$y_{i0} = \left(\begin{bmatrix} \rho_i & 1 \end{bmatrix} \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_0^2 \end{bmatrix}^{-1} \begin{bmatrix} \rho_i \\ 1 \end{bmatrix}\right)^{-1} \left(\begin{bmatrix} \rho_i & 1 \end{bmatrix} \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_0^2 \end{bmatrix}^{-1} \begin{bmatrix} y_{i1} - x_{i1}'\beta_i + u_{i1} \\ \bar{y}_0 \end{bmatrix}\right) \tag{1.32}$$

for each $i = 1, 2, \ldots, N$.

1.4.6 Efficiency Measurement

Once the sequence of the conditional posterior technical inefficiencies are obtained from the Gibbs Sampling iterations:

$$\boldsymbol{u}_i^1, \boldsymbol{u}_i^2, \boldsymbol{u}_i^3, ..., \boldsymbol{u}_i^S \tag{1.33}$$

where S is the number of MCMC iterations, a common approach to obtain estimates of a firm's inefficiency level is to use the average of the draws, as:

$$\hat{\boldsymbol{u}}_i = S^{-1} \sum_{s=1}^{S} \boldsymbol{u}_i^s \tag{1.34}$$

Hence, the corresponding technical efficiency scores can be obtained by utilizing the definition of technical efficiency, as:

$$\hat{\boldsymbol{TE}}_i = exp(-\hat{\boldsymbol{u}}_i) \tag{1.35}$$

where $\hat{\boldsymbol{TE}}_i$ is a $T \times 1$ vector of the estimated efficiencies of firm i.

1.5 Performance of the model using artificial data

1.5.1 Data Generating Process

In this section, we present an experiment in order to evaluate the finite sample performance of the proposed MCMC algorithm. In particular, for the purpose of this experiment, we estimate a dynamic panel random coefficients production frontier model with a single input x_{it} according to the following Data Generated Process (DGP):

$$y_{it} = a_i + \rho_i y_{it-1} + \beta_i x_{it} + v_{it} - u_{it}$$
$$v_{it} \sim \mathcal{N}(0, \sigma_v^2)$$
$$u_{it} \sim \mathcal{N}^+(0, \sigma_u^2)$$

The random coefficients a_i, ρ_i and β_i are generated using the multivariate normal distribution structure as:

$$\begin{bmatrix} a_i \\ \rho_i \\ \beta_i \end{bmatrix} \sim \boldsymbol{\mathcal{N}} \left(\begin{bmatrix} \bar{a} \\ \bar{\rho} \\ \bar{\beta} \end{bmatrix}, \begin{bmatrix} \sigma_a^2 & 0 & 0 \\ 0 & \sigma_\rho^2 & 0 \\ 0 & 0 & \sigma_\beta^2 \end{bmatrix} \right)$$

where for simplicity we assume a diagonal variance-covariance matrix of the parameters which implies that the parameters are generated independently from each other. We perform this experiment for different settings, where we evaluate our model for the different number of cross-sectional units $N = \{50, 100\}$, different time periods $T = \{10, 20\}$ and different values of the autoregressive parameters $\bar{\rho} = \{0.2, 0.5, 0.8\}$.

The input variable x_{it} is generated from a $\mathcal{N}(0, 1)$ and can be seen as a transformed variable that reflects deviations of the original input, let's say X_{it}, from the average $\bar{X}$, in logarithm form[7]. In order to secure that the autoregressive parameters ρ_i lie on the interval $[0, 1]$, we pick $\sigma_\rho = 0.05$ for the cases where $\bar{\rho} = \{0.2, 0.8\}$. For the scenario where $\bar{\rho} = 0.5$ we choose $\sigma_\rho = 0.1$, since it is ensured that the generated autoregressive parameters will lie on the required interval[8]. Last, for the β_i's coefficients, for all scenarios we use $\bar{\beta} = 1$ and $\sigma_\beta = 0.1$. In general, we believe that the choice of all the above parameters is realistic and can describe real empirical specifications.

In addition, for all the different experimental settings, in order to ensure the existence of

[7] Such a parameterization is very common in many empirical applications, especially in cases where macroeconomic variables are incorporated in the analysis.

[8] Another approach, where we can ensure the stationarity of the process in our DGP is to follow closely Hsiao et al. (1999) and for the data generating process of the autoregressive parameter to use the truncated normal distribution.

the above effects in our generated data, for each unit i we generate $m + T$ time periods, where the first m time periods are dropped out from our analysis. We set $m = 10$.

Furthermore, we pay particular attention to the signal-to-noise ratio, defined as $\mathcal{F} = \sigma_u/\sigma_v$, since, from the stochastic frontier literature, it is well known that small values of $\mathcal{F}$ create identification issues regarding the inefficiency estimates. The rationale here is that as $\sigma_v \to \infty$ and consequently $\mathcal{F} \to 0$, the distribution of the common error term will dominate the probability density function of the inefficiencies u_{it}, and as a result, the identification of technical inefficiency will not be possible. On the other hand, as $\sigma_u \to \infty$ there will be enough evidence to identify the latent u_{it} term. In Table 1.1, we present the different experiment settings.

Table 1.1: Experiment Settings

Experiment Setting (I)	$\sigma_v = 0.1$	$\sigma_u = 0.2$	$\mathcal{F} = 2$	$N = \{50, 100\}$	$T = \{10, 20\}$
Experiment Setting (II)	$\sigma_v = 0.1$	$\sigma_u = 0.1$	$\mathcal{F} = 1$	$N = \{50, 100\}$	$T = \{10, 20\}$
Experiment Setting (III)	$\sigma_v = 0.05$	$\sigma_u = 0.15$	$\mathcal{F} = 3$	$N = \{50, 100\}$	$T = \{10, 20\}$
Experiment Setting (IV)	$\sigma_v = 0.01$	$\sigma_u = 0.05$	$\mathcal{F} = 5$	$N = \{50, 100\}$	$T = \{10, 20\}$

1.5.2 Prior Specifications

In this subsection, we illustrate our choice for the prior hyperparameters. More specifically, for the mean value and the variance-covariance matrix of the parameters and the set $\boldsymbol{\mu}_0 = \mathbf{0}_{3\times1}$ and $\boldsymbol{\Lambda}_0 = 10^3 \times \boldsymbol{I}_{3\times3}$. This particular choice generates a pretty diffuse normal prior distribution; hence, our prior specification cannot dominate the likelihood function. For the variances of the error term σ_v^2 and the variance of the inefficiency term σ_u^2, we set $a_0 = a_1 = 10^{-6}$ and $\gamma_0 = \gamma_1 = 10^{-6}$. As discussed above, these hyperparameters generate the usual non-informative Jeffrey's prior distribution. Last, for the variance-covariance matrix, we set $v_0 = 0$ and $\boldsymbol{\Psi}_0 = 10^{-6} \times \boldsymbol{I}_{3\times3}$. These hyperparameters create vague prior probability densities.

1.5.3 Results

In Tables 1.2-1.5 we present the posterior estimates of our experiments. In particular, for the different number of firms N and time periods T, we report the conditional posterior average and the posterior standard deviation for all the unknown parameters. In addition, for each experiment, we report the Pearson correlation between the true (the generated) and the estimated inefficiencies. The Bayesian MCMC is based on 10,000 iterations from which the first 5,000 posterior draws are discarded from our analysis in order to eliminate potential effects of the initial values.

In general, we see that for all the different experiment designs, the corresponding posterior densities are distributed around the true values of the parameters. In particular, we see that for all the generated posterior densities, the true parameter values belong to 95% credible interval. This illustrates that our proposed Bayesian hierarchical model with data augmentation is able to estimate consistently all the structural parameters of interest.

In addition, we see that as the signal-to-noise ratio $\mathcal{F}$ increases, the Pearson correlation between the real and the estimated inefficiency scores tends to unity. More specifically, for the Experiment Setting (I) presented in Table 1.2, where the $\mathcal{F} = 2$ we see that the correlation coefficient is around 0.70 and increases to 0.76, as T increases from 10 to 20. On the other hand, as λ increases to 3 and 5, as presented in Experiment Setting (II) and (III) in Tables 1.4 and 1.5, respectively, we observe that the correlation coefficient increases from around 0.78 and 0.85, for $T = 10$ and $T = 20$, to 0.84 and 0.91, for $T = 10$ and $T = 20$, respectively. On the contrary, as we can see from Table 1.3 where the signal-to-noise ratio $\mathcal{F} = 1$, the correlation between the true and estimated inefficiencies drop to range between 0.45 and 0.53, for the two different time periods $T = \{10, 20\}$. These statistical properties of the proposed model are a-priori expected and indicate that our proposed model is able to identify the latent inefficiency term, as long as the inefficiency signal is adequate to draw posterior inference.

Table 1.2: Posterior Estimates for the Experiment Setting (I)

Panel A

		$N = 50$				$N = 100$			
		$T = 10$		$T = 20$		$T = 10$		$T = 20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	1.0240	0.0243	1.0045	0.0209	1.0227	0.0168	1.0001	0.0138
$\bar{\rho}$	0.20	0.2130	0.0126	0.1905	0.0110	0.1987	0.0085	0.2094	0.0073
$\bar{\beta}$	1.00	1.0027	0.0165	1.0050	0.0168	1.0095	0.0122	1.0111	0.0111
σ_a	0.10	0.1050	0.0128	0.1007	0.0116	0.0941	0.0092	0.0984	0.0081
σ_ρ	0.05	0.0722	0.0084	0.0692	0.0076	0.0640	0.0059	0.0637	0.0052
σ_β	0.10	0.1053	0.0116	0.1125	0.0122	0.1091	0.0093	0.1031	0.0081
σ_v	0.10	0.0841	0.0146	0.1032	0.0099	0.0909	0.0102	0.0952	0.0064
σ_u	0.20	0.2130	0.0212	0.2004	0.0171	0.2118	0.0151	0.2051	0.0106
$cor(u_{it}; \hat{u}_{it})$		0.7056		0.7481		0.7168		0.7650	

Panel B

		$N = 50$				$N = 100$			
		$T = 10$		$T = 20$		$T = 10$		$T = 20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	1.0176	0.0242	1.0043	0.0208	1.0248	0.0185	0.9971	0.0142
$\bar{\rho}$	0.50	0.5195	0.0169	0.4731	0.0151	0.4961	0.0115	0.5098	0.0111
$\bar{\beta}$	1.00	1.0036	0.0168	1.0055	0.0171	1.0099	0.0121	1.0110	0.0111
σ_a	0.10	0.1081	0.0150	0.1024	0.0124	0.0974	0.0108	0.1030	0.0091
σ_ρ	0.10	0.1079	0.0116	0.1006	0.0103	0.1050	0.0082	0.1055	0.0076
σ_β	0.10	0.1067	0.0123	0.1116	0.0122	0.1077	0.0090	0.1036	0.0081
σ_v	0.10	0.0805	0.0119	0.1024	0.0092	0.0863	0.0105	0.0958	0.0061
σ_u	0.20	0.2181	0.0173	0.2029	0.0151	0.2178	0.0148	0.2040	0.0100
$cor(u_{it}; \hat{u}_{it})$		0.7098		0.7469		0.7078		0.7639	

Panel C

		$N = 50$				$N = 100$			
		$T = 10$		$T = 20$		$T = 10$		$T = 20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	1.0016	0.0311	0.9899	0.0284	0.9968	0.0240	0.9992	0.0177
$\bar{\rho}$	0.80	0.8128	0.0112	0.7869	0.0096	0.8021	0.0074	0.8047	0.0066
$\bar{\beta}$	0.10	1.0059	0.0166	1.0071	0.0167	1.0102	0.0117	1.0116	0.0110
σ_a	0.10	0.1118	0.0194	0.1046	0.0153	0.0966	0.0143	0.1077	0.0118
σ_ρ	0.05	0.0671	0.0070	0.0636	0.0064	0.0605	0.0046	0.0593	0.0043
σ_β	0.10	0.1042	0.0115	0.1107	0.0118	0.1071	0.0091	0.1037	0.0079
σ_v	0.10	0.0817	0.0158	0.1111	0.0110	0.0937	0.0103	0.0929	0.0069
σ_u	0.20	0.2163	0.0222	0.1881	0.0214	0.2088	0.0153	0.2085	0.0108
$cor(u_{it}; \hat{u}_{it})$		0.7134		0.7495		0.7176		0.7668	

Table 1.3: Posterior Estimates for the Experiment Setting (II)

Panel A

		$N=50$				$N=100$			
		$T=10$		$T=20$		$T=10$		$T=20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	1.0342	0.0285	1.0004	0.0232	0.9802	0.0304	0.9944	0.0162
$\bar{\rho}$	0.20	0.2098	0.0114	0.1890	0.0104	0.1980	0.0074	0.2083	0.0067
$\bar{\beta}$	1.00	1.0002	0.0164	1.0043	0.0165	1.0064	0.0117	1.0116	0.0108
σ_a	0.10	0.1009	0.0120	0.1001	0.0111	0.0969	0.0088	0.0999	0.0076
σ_ρ	0.05	0.0703	0.0078	0.0666	0.0070	0.0613	0.0054	0.0618	0.0047
σ_β	0.10	0.1055	0.0112	0.1110	0.0116	0.1082	0.0086	0.1031	0.0078
σ_v	0.10	0.0895	0.0127	0.1043	0.0073	0.1097	0.0088	0.0981	0.0056
σ_u	0.10	0.1195	0.0297	0.0968	0.0218	0.0613	0.0347	0.0968	0.0154
$cor(u_{it};\hat{u}_{it})$		0.4591		0.5077		0.4888		0.5235	

Panel B

		$N=50$				$N=100$			
		$T=10$		$T=20$		$T=10$		$T=20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	1.0210	0.0381	0.9483	0.0355	0.9807	0.0332	0.9935	0.0181
$\bar{\rho}$	0.50	0.5147	0.0166	0.4718	0.0146	0.4955	0.0114	0.5090	0.0106
$\bar{\beta}$	1.00	1.0008	0.0163	1.0039	0.0164	1.0072	0.0116	1.0118	0.0110
σ_a	0.10	0.1039	0.0136	0.1020	0.0120	0.0999	0.0102	0.1048	0.0088
σ_ρ	0.10	0.1079	0.0113	0.1001	0.0102	0.1043	0.0078	0.1046	0.0075
σ_β	0.10	0.1058	0.0113	0.1107	0.0116	0.1075	0.0085	0.1034	0.0077
σ_v	0.10	0.0906	0.0172	0.1159	0.0084	0.1082	0.0106	0.0977	0.0058
σ_u	0.10	0.1114	0.0426	0.0349	0.0396	0.0655	0.0376	0.0978	0.0171
$cor(u_{it};\hat{u}_{it})$		0.4640		0.5084		0.4836		0.5232	

Panel C

		$N=50$				$N=100$			
		$T=10$		$T=20$		$T=10$		$T=20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	1.0371	0.0298	0.9855	0.0275	0.9474	0.0314	0.9988	0.0186
$\bar{\rho}$	0.80	0.8087	0.0106	0.7860	0.0096	0.8010	0.0069	0.8044	0.0062
$\bar{\beta}$	1.00	1.0006	0.0160	1.0046	0.0163	1.0080	0.0117	1.0119	0.0107
σ_a	0.10	0.1080	0.0183	0.1066	0.0143	0.0989	0.0142	0.1131	0.0116
σ_ρ	0.05	0.0668	0.0067	0.0632	0.0063	0.0600	0.0044	0.0587	0.0042
σ_β	0.10	0.1047	0.0113	0.1097	0.0112	0.1073	0.0083	0.1030	0.0076
σ_v	0.10	0.0783	0.0122	0.1083	0.0065	0.1126	0.0071	0.0953	0.0051
σ_u	0.10	0.1412	0.0210	0.0839	0.0231	0.0475	0.0306	0.1053	0.0128
$cor(u_{it};\hat{u}_{it})$		0.4679		0.5067		0.4890		0.5286	

Table 1.4: Posterior Estimates for the Experiment Setting (III)

Panel A

		$N = 50$				$N = 100$			
		$T = 10$		$T = 20$		$T = 10$		$T = 20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	1.0198	0.0176	1.0044	0.0168	1.0146	0.0130	0.9997	0.0110
$\bar{\rho}$	0.20	0.2112	0.0111	0.1883	0.0099	0.1978	0.0072	0.2067	0.0066
$\bar{\beta}$	1.00	0.9973	0.0154	1.0036	0.0162	1.0067	0.0114	1.0090	0.0105
σ_a	0.10	0.1009	0.0114	0.1027	0.0107	0.0941	0.0080	0.0977	0.0072
σ_ρ	0.05	0.0678	0.0075	0.0654	0.0068	0.0625	0.0052	0.0621	0.0047
σ_β	0.10	0.1032	0.0106	0.1112	0.0114	0.1074	0.0081	0.1012	0.0075
σ_v	0.05	0.0381	0.0086	0.0528	0.0058	0.0425	0.0079	0.0489	0.0040
σ_u	0.15	0.1560	0.0105	0.1488	0.0082	0.1551	0.0084	0.1514	0.0054
$cor(u_{it}; \hat{u}_{it})$		0.7886		0.8422		0.7901		0.8516	

Panel B

		$N = 50$				$N = 100$			
		$T = 10$		$T = 20$		$T = 10$		$T = 20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	1.0160	0.0200	1.0028	0.0168	1.0131	0.0131	1.0004	0.0117
$\bar{\rho}$	0.50	0.5166	0.0160	0.4709	0.0144	0.4944	0.0110	0.5079	0.0108
$\bar{\beta}$	1.00	0.9975	0.0157	1.0033	0.0162	1.0067	0.0113	1.0097	0.0103
σ_a	0.10	0.1007	0.0119	0.1034	0.0115	0.0948	0.0089	0.0999	0.0080
σ_ρ	0.10	0.1071	0.0111	0.1000	0.0100	0.1058	0.0077	0.1058	0.0076
σ_β	0.10	0.1039	0.0108	0.1107	0.0112	0.1067	0.0081	0.1015	0.0073
σ_v	0.05	0.0317	0.0111	0.0522	0.0053	0.0404	0.0067	0.0471	0.0043
σ_u	0.10	0.1611	0.0117	0.1498	0.0074	0.1570	0.0081	0.1535	0.0058
$cor(u_{it}; \hat{u}_{it})$		0.7880		0.8417		0.7862		0.8504	

Panel C

		$N = 50$				$N = 100$			
		$T = 10$		$T = 20$		$T = 10$		$T = 20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	0.9999	0.0239	0.9992	0.0194	0.9977	0.0178	0.9960	0.0137
$\bar{\rho}$	0.80	0.8125	0.0100	0.7857	0.0092	0.7998	0.0066	0.8042	0.0062
$\bar{\beta}$	1.00	0.9991	0.0153	1.0045	0.0162	1.0070	0.0113	1.0098	0.0104
σ_a	0.10	0.1082	0.0158	0.1042	0.0131	0.0942	0.0116	0.1022	0.0096
σ_ρ	0.05	0.0658	0.0066	0.0625	0.0061	0.0597	0.0045	0.0594	0.0042
σ_β	0.10	0.1029	0.0106	0.1097	0.0110	0.1063	0.0081	0.1011	0.0073
σ_v	0.05	0.0275	0.0124	0.0543	0.0055	0.0421	0.0086	0.0473	0.0039
σ_u	0.15	0.1633	0.0100	0.1478	0.0080	0.1558	0.0095	0.1532	0.0055
$cor(u_{it}; \hat{u}_{it})$		0.7969		0.8445		0.7997		0.8523	

Table 1.5: Posterior Estimates for the Experiment Setting (IV)

Panel A									
		$N = 50$				$N = 100$			
		$T = 10$		$T = 20$		$T = 10$		$T = 20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	1.0214	0.0143	1.0054	0.0146	1.0091	0.0101	1.0039	0.0097
$\bar{\rho}$	0.20	0.2070	0.0094	0.1846	0.0088	0.1958	0.0059	0.2037	0.0060
$\bar{\beta}$	1.00	0.9910	0.0146	1.0005	0.0153	1.0037	0.0107	1.0083	0.0101
σ_a	0.10	0.0970	0.0096	0.1030	0.0102	0.0987	0.0071	0.0963	0.0066
σ_ρ	0.05	0.0649	0.0064	0.0619	0.0059	0.0593	0.0042	0.0597	0.0042
σ_β	0.10	0.1020	0.0098	0.1082	0.0105	0.1061	0.0074	0.0996	0.0071
σ_v	0.01	0.0051	0.0042	0.0123	0.0015	0.0044	0.0023	0.0092	0.0012
σ_u	0.05	0.0505	0.0028	0.0484	0.0019	0.0515	0.0018	0.0510	0.0013
$cor(u_{it}; \hat{u}_{it})$		0.8323		0.9105		0.8313		0.9113	

Panel B									
		$N = 50$				$N = 100$			
		$T = 10$		$T = 20$		$T = 10$		$T = 20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	1.0193	0.0143	1.0048	0.0149	1.0062	0.0103	1.0035	0.0099
$\bar{\rho}$	0.50	0.5105	0.0151	0.4676	0.0142	0.4929	0.0104	0.5056	0.0106
$\bar{\beta}$	1.00	0.9914	0.0144	1.0011	0.0155	1.0037	0.0107	1.0087	0.0100
σ_a	0.10	0.0964	0.0099	0.1026	0.0100	0.0981	0.0071	0.0970	0.0070
σ_ρ	0.10	0.1060	0.0102	0.0988	0.0098	0.1045	0.0073	0.1051	0.0073
σ_β	0.10	0.1018	0.0100	0.1088	0.0107	0.1060	0.0074	0.0994	0.0070
σ_v	0.01	0.0069	0.0020	0.0117	0.0016	0.0073	0.0021	0.0092	0.0013
σ_u	0.05	0.0501	0.0024	0.0489	0.0021	0.0502	0.0020	0.0510	0.0014
$cor(u_{it}; \hat{u}_{it})$		0.8387		0.9104		0.8371		0.9112	

Panel C									
		$N = 50$				$N = 100$			
		$T = 10$		$T = 20$		$T = 10$		$T = 20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	1.0135	0.0151	1.0051	0.0150	1.0011	0.0110	1.0009	0.0101
$\bar{\rho}$	0.80	0.8068	0.0093	0.7837	0.0088	0.7970	0.0061	0.8030	0.0058
$\bar{\beta}$	1.00	0.9922	0.0145	1.0006	0.0153	1.0035	0.0106	1.0084	0.0100
σ_a	0.10	0.0951	0.0107	0.1028	0.0104	0.0978	0.0080	0.0969	0.0072
σ_ρ	0.05	0.0648	0.0063	0.0617	0.0060	0.0592	0.0042	0.0591	0.0042
σ_β	0.10	0.1023	0.0100	0.1080	0.0105	0.1056	0.0074	0.0996	0.0070
σ_v	0.01	0.0046	0.0026	0.0117	0.0016	0.0078	0.0018	0.0095	0.0012
σ_u	0.05	0.0515	0.0026	0.0492	0.0021	0.0503	0.0019	0.0507	0.0014
$cor(u_{it}; \hat{u}_{it})$		0.8344		0.9104		0.8477		0.9091	

1.6 Empirical Application

1.6.1 Data

In this empirical application, we employ the proposed model in the same empirical application presented in Bhattacharyya (2012), where the author employs the two-stage approach in order to estimate the technical efficiency scores of 28 Egyptian manufacturing sectors during the financial years 1987/1988 to 1995/1996 (9 years). The data used in the paper are coming directly from the study initially performed by Getachew and Sickles (2007) where the authors analyze the cost performance of the Egyptian private manufacturing sectors[9]. In particular, Bhattacharyya (2012) follow a different approach and investigate the production performance of the manufacturing sectors by employing the dynamic panel production frontier model. As mentioned by the author, during the period of the early 90s, the Egyptian government undertook privatization reforms which generated many new employment opportunities for unskilled and semi-skilled labour, and as a result, the different sectors were affected by the sluggish adjustment of the new workers.

The dependent variable is the Output Quantity Index measured as the total revenue deflated by the relevant price indices. For the inputs used in the process, we have data on Capital, Labour, Energy, and the corresponding Materials used in each sector. Further details and summary statistics of this dataset can be found in Getachew and Sickles (2007) and Bhattacharyya (2012).

The objective of this empirical application is twofold. First, we want to illustrate the differences in technical efficiency estimates between the static and the dynamic stochastic frontier specification. Secondly, we want to compare the results obtained from the proposed Bayesian hierarchical dynamic panel model with those obtained from the two-stage approach. The cross-examination of the two approaches is based on the estimates of (i) the autoregressive parameter, (ii) the short-run input elasticities, and (iii) the level of the estimated technical efficiency scores. Additionally, using the proposed DRC-SFM, we present the estimated technological progress, efficiency change and Total Factor Productivity.

[9] The dataset is available by the Journal of Applied Econometrics Data Archive.

1.6.2 The Empirical Model

For the empirical model we follow closely Bhattacharyya (2012) and adopt a simple Cobb-Douglas production function[10]. The empirical model under the dynamic specification will be:

$$y_{it} = \rho_i \, y_{it-1} + \beta_{Ki} \, k_{it} + \beta_{Li} \, l_{it} + \beta_{Ei} \, e_{it} + \beta_{Mi} \, m_{it} + a_i + \boldsymbol{\gamma}_i' \boldsymbol{f}_t + v_{it} - u_{it} \tag{1.36}$$

where k_{it}, l_{it}, e_{it} and m_{it} are the capital, labour, energy and materials used in each sector i at time t. In the analysis, all the variables are used in natural logarithm form. In addition, to account for the sector-specific technological progress or regress, we include the vector $\boldsymbol{f}_t$ which contains the linear and quadratic time trend. Last, as in the previous sections for the measurement errors we assume $v_{it} \sim \mathcal{N}(0, \sigma_v^2)$ and for the non-negative inefficiency term we have $u_{it} \sim \mathcal{N}^+(0, \sigma_u^2)$.

The above dynamic specification enables us to obtain both the average short-run and the long-run input elasticities as:

$$\text{Short-Run Effects:} \quad \frac{\partial y_{it}}{\partial k_{it}} = \bar{\beta}_K \quad , \quad \frac{\partial y_{it}}{\partial l_{it}} = \bar{\beta}_L \quad , \quad \frac{\partial y_{it}}{\partial e_{it}} = \bar{\beta}_E \quad , \quad \frac{\partial y_{it}}{\partial m_{it}} = \bar{\beta}_M$$

$$\text{Long-Run Effects:} \quad \frac{\partial y_{it}}{\partial k_{it}} = \frac{\bar{\beta}_K}{1-\bar{\rho}} \quad , \quad \frac{\partial y_{it}}{\partial l_{it}} = \frac{\bar{\beta}_L}{1-\bar{\rho}} \quad , \quad \frac{\partial y_{it}}{\partial e_{it}} = \frac{\bar{\beta}_E}{1-\bar{\rho}} \quad , \quad \frac{\partial y_{it}}{\partial m_{it}} = \frac{\bar{\beta}_M}{1-\bar{\rho}}$$

Moreover, we can distinguish between the technical inefficiency that arises from the fact that sectors do not utilize perfectly their input variables, and the output fall due to the firm's short-run inability to operate at full capacity. Specifically, we have:

$$u_{it} = \text{technical inefficiency obtained from the model} \quad , \quad u_{it}^{adj} = \frac{u_{it}}{1-\bar{\rho}} - u_{it}$$

where u_{it}^{adj} illustrates the contribution of adjustment costs to the overall output fall.

As discussed extensively in the previous sections, under the assumption that output adjusts instantaneously to the long-run equilibrium, viz. $\rho_i = 0$ for each $i = 1, 2, \ldots, N$, the proposed specification nests previous panel stochastic frontier models. For comparison reasons, we also estimate a panel stochastic frontier model under a static specification that has dominated the empirical literature. In particular, we present estimates of the True Random Effects (TRE) panel stochastic frontier model with fixed coefficients as proposed by Greene (2005a,b), where we can control for the time-invariant heterogeneity effects across the sectors. As in the case of the dynamic specification, in order for our results to be directly comparable, for the TRE-SF model we adopt a hierarchical structure where the estimation inference is based on Bayesian

[10] We performed the same analysis using the translog production function and we found that the technical efficiency estimates remain fundamentally the same. For this reason, we chose to report only the results obtained from the Cobb-Douglas specification.

MCMC techniques[11]. In Table 1.6, we summarize the two different model specifications used in this study.

Table 1.6: Bayesian Hierarchical Empirical Models

Static Specification		
(Assumption under Perfect Adjustment)		
M1	**BH-TRE-SFM**	$y_{it} = x'_{it}\beta + a_i + v_{it} - u_{it}$
		$v_{it} \sim \mathcal{N}(0, \sigma_v^2)$, $u_{it} \sim \mathcal{N}(0, \sigma_u^2)$
		$a_i \sim \mathcal{N}(0, \sigma_a^2)$
Dynamic Specification		
(Assumption under Partial Adjustment)		
M2	**BH-DRC-SFM**	$y_{it} = \rho_i y_{it-1} + x'_{it}\beta_i + v_{it} - u_{it}$
		$v_{it} \sim \mathcal{N}(0, \sigma_v^2)$, $u_{it} \sim \mathcal{N}(0, \sigma_u^2)$
		$[\rho_i, \beta_i]' \sim \mathcal{N}\left([\bar{\rho}, \bar{\beta}]', \Sigma\right)$

1.6.3 Prior Specifications

Before we proceed with the empirical results we present the prior elicitation for our dynamic random coefficient model. For the variance of the error term σ_v^2 we have $a_0 = a_1 = 10^{-2}$. For the variance of the inefficiency term σ_u^2 we specify $\gamma_0 = 3$ and $\gamma_1 = 1$ which leads to a highly non-informative Inverse-Gamma distribution. For the random coefficients we have $\mu_0 = 0 \times \mathbf{1}_{J\times 1}$ and $\Lambda_0 = 100 \times \boldsymbol{I}_{J\times J}$. Last, for the variance-covariance matrix $\boldsymbol{\Omega}$ we choose $v_0 = 0$ and $\boldsymbol{\Psi}_0 = 10^{-6} \times \boldsymbol{I}_{J\times J}$. In addition, for the BH-TRE-SFM model, where $a_i \sim \mathcal{N}(0, \sigma_a^2)$, for the variance σ_a^2 the prior specification is assigned as $\sigma_a^2 | q_0, q_1 \sim \mathcal{IG}(q_0, q_1)$, with $q_0 = q_1 = 10^{-2}$. Last, for the slope parameters for the BH-TRE-SFM we use the uninformative flat prior, viz. $\pi(\beta) \propto 1$.

1.6.4 Empirical Results

In Table 1.7, we report the estimated parameters using both the SYS-GMM approach and the proposed Bayesian hierarchical modeling. More specifically, in the first two columns, we present the GMM estimates for the static and the dynamic specification, as presented in the original paper. In the third and fourth column, we report posterior averages and posterior standard deviations of the parameters using the BH-TRE-SFM and BH-DRC-SFM, respectively. For the two hierarchical models, all Gibbs Sampling computations are conducted using 30,000 iterations from which the first 10,000 draws are discarded from our analysis to mitigate any initial values

[11] In Appendix A, we present the Bayesian posterior analysis of the TRE-SFM.

Table 1.7: Posterior estimates of parameters and technical efficiency scores

Parameters	Description	SYS-GMM		Bayesian Hierarchical	
		Static	Dynamic	TRE-SFM	DRC-SFM
Frontier Parameters					
$\bar{a}$	Constant	0.803***	0.33	0.2597	0.2619
		(0.15)	(0.29)	(0.0731)	(0.0837)
$\bar{\rho}$	Lagged Log-Output	-	0.16***	-	0.1565
			(0.06)		(0.0248)
$\bar{\beta}_K$	Capital Elas.	0.014	0.02	0.0228	0.0214
		(0.01)	(0.05)	(0.0161)	(0.0183)
$\bar{\beta}_L$	Labour Elas.	0.123***	0.22***	0.1439	0.1506
		(0.04)	(0.09)	(0.0440)	(0.0502)
$\bar{\beta}_E$	Energy Elas.	0.044***	0.04	0.0322	0.0294
		(0.02)	(0.05)	(0.0241)	(0.0245)
$\bar{\beta}_M$	Material Elas.	0.833**	0.65***	0.7948	0.6448
		(0.03)	(0.09)	(0.0439)	(0.0485)
σ_a	Heterogeneity Std.			0.1853	
				(0.0363)	
σ_v	Noise Std.			0.1438	0.0537
				(0.0142)	(0.0129)
σ_u	Inefficiency Std.			0.2552	0.2495
				(0.0222)	(0.0183)
Technical Efficiency Estimates					
Median		70.0%	74.5%	83.5%	84.5%
Mean		68.3%	74.7%	82.5%	82.9%
Max		100.0%	100.0%	96.4%	96.6%
Min		59.6%	54.2%	47.6%	53.6%

Notes*: The results using the SYS-GMM technique are reported as presented by Bhattacharyya (2012). For the Bayesian hierarchical specifications, we report the posterior average and posterior standard deviation of the parameters. The MCMC algorithm is based on 30,000 iterations from which the first 10,000 samples are discarded. For the SYS-GMM the time effects are captured by using $T-1$ time dummy variables. We do not report those estimates to save space.

effect. Therefore, the empirical inference is based on the remaining 20,000 posterior sample draws.

First, from the static and the dynamic Bayesian hierarchical specifications presented in the third and fourth column of Table 1.7, we see that dynamic specification generates higher efficiency estimates. This illustrates the importance of controlling the production dynamics when the efficiency evaluation is particularly important. Regarding the potential technical efficiency improvements, the results indicate that the manufacturing sectors were able to increase their production output and reduce their average production cost on average by 20%[12].

[12]The potential output increase can be calculated as 1/0.83=1.20, where 0.83 is the average technical efficiency

Comparing the efficiency estimates obtained from the hierarchical models with those obtained from the two-stage approach, we see that the relative efficiency estimates using the Schmidt and Sickles (1984) method will underestimate the real technical efficiency level. In particular, the average efficiency score obtained by the SYS-GMM method, presented in the second column of the Table 1.7 is 74.7%, around 8% lower than the estimate provided by the BH-DRC-SFM. These results were anticipated, since as we illustrated in the previous sections, the relative comparison method is not able to separate the sector-specific unobserved effects from the inefficiency effects, and as a result, the generated efficiency scores will be distorted.

Regarding the mean autoregressive parameter $\bar{\rho}$, we see that the posterior average is 0.1565 with a posterior standard deviation of 0.0248, which implies that the data do not support the assumption of instantaneous production adjustment. More specifically, the posterior average implies an average production catch-up rate of 84.4%. This result is numerically very similar to the System-GMM approach presented in the second column of Table 1.7. Moreover, in Figure 1.1, we illustrate the marginal posterior distribution of each autoregressive parameter ρ_i. From the graph, it is clear that there is evidence that the speed of adjustment is quite heterogeneous among the different sectors, and each sector is facing a different speed of adjustment towards the long-run equilibrium.

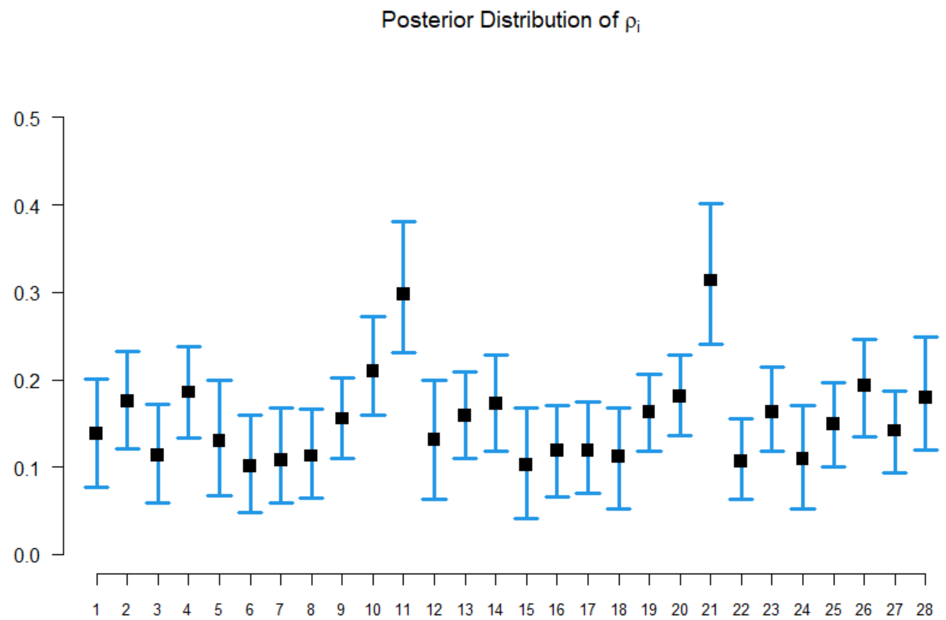


Figure 1.1: 68% Credible Interval of the autoregressive parameters ρ_i

In addition, in Table 1.8, we report the corresponding annual adjustment rate of pro-

score obtained from the BH-DRC-SFM.

duction for each sector. Overall, the adjustment rate is relatively high, indicating that the adjustment frictions generated by the market liberalization, such as labour mobility, had a mild impact on the manufacturing sector. As Getachew and Sickles (2007) highlighted, during that period the Egyptian economy faced high unemployment rates and one plausible explanation is that labour could be replaced more easily. Moreover, another factor that could have contributed to low adjustment frictions is the absence of sufficient labour market regulations making the labour adjustment frictions less costly (e.g. absence of hiring and firing costs, elastic labour arrangements, etc.).

Table 1.8: Period required for production full adjustment

Industrial Activity	Annual Adjustment Rate	Industrial Activity	Annual Adjustment Rate
Food Manufacturing	0.861	Other petroleum and coal	0.898
Other Food Manufacturing	0.825	Manufacture of rubber products	0.881
Beverage and liquor	0.886	Manufacture of plastic products	0.881
Tobacco	0.815	Manufacture of pottery and china	0.888
Manufacture of textile	0.870	Manufacture of glass and glass products	0.837
Manufacture of wearing apparels	0.899	Manufacture of other non metallic products	0.819
Manufacture of leather products	0.892	Iron and steel basic industries	0.686
Manufacture of footwear	0.887	Non-ferrous basic industries	0.894
Manufacture of wood products	0.845	Manufacture of fabricated metal products	0.837
Manufacture of furniture & fixture	0.790	Manufacture of machinery except electrical	0.891
Manufacture of paper products	0.702	Manufacture of electrical machinery	0.851
Printing and publishing industries	0.869	Manufacture of transport equipment	0.807
Manufacture of industrial chemicals	0.842	Manufacture of professional equipment	0.859
Manufacture of other chemical products	0.827	Other manufacture industries	0.820

Notes*: The Annual Adjustment Rate is calculated as $1 - \bar{\rho}_i$, where $\bar{\rho}_i$ is the posterior average of the sector specif autoregressive parameter.

Besides the efficiency scores, from the BH-DRC-SFM, we are able to obtain the implied overall output loss. We see that the median and the average output loss are 20.0% and 23.2%, respectively. These posterior estimates are approximately 3.0% higher than the technical inefficiency scores obtained from the Bayesian DRC specification. In economic terms, this indicates that the short-run input adjustment frictions, contribute to the overall output loss on average by 3%. In Figure 1.2, we illustrate the difference between technical inefficiency and overall output fall. As a result, it is evident that ignoring for the dynamic structure and the intertemporal optimization problem, static specifications can lead to bias technical efficiency estimates.

Regarding the Return to scales, from the BH-DRC-SFM we see that the posterior average of the short-run RTS is 0.8462. This implies that in the short-run if the firms decide to scale up their production and increase their inputs by 10%, the total cost will increase by 11.82%. This suggests, that firms will face a 1.82% additional cost, which should be attributed to the short-

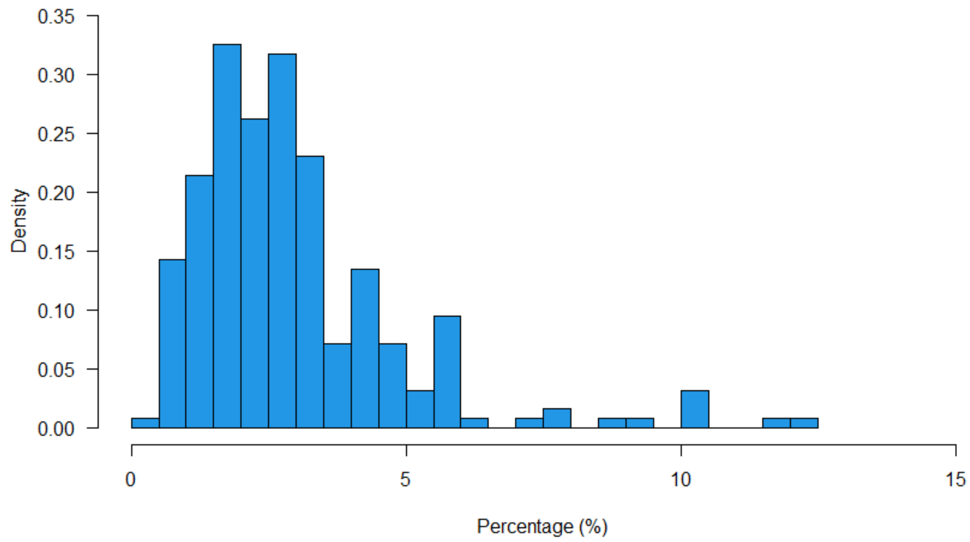


Figure 1.2: Output fall related to partial input contribution.

Table 1.9: Long-Run Posterior Estimates

Input Elasticities and Returns to Scales					
	Capital Elas.	Labor Elas.	Energy Elas.	Material Elas.	RTS
Post. Mean	0.0255	0.1784	0.0349	0.7642	1.0030
Post. Std.	(0.0219)	(0.0594)	(0.0349)	(0.0511)	(0.0456)
Overall Output Loss					
	Median	Mean	Max	Min	
	20.01%	23.22%	73.95%	4.20%	

Notes*: This table reports the posterior average and posterior standard deviation of the long-run input elasticities, the RTS, and the Technical Efficiency scores. As in Table 1.7, the estimates are based on 30,000 MCMC iterations from which the first 10,000 are discarded.

run adjustment costs. Moreover, in Table 1.9, we present the posterior averages and standard deviations of the long-run input elasticities, the RTS, and the overall output-loss. The posterior average of the long-run RTS is 1.003, indicating that our model provides estimates consistent with the economic theory.

Last, comparing the technical efficiency scores obtained from the dynamic and the mis-specified static Bayesian hierarchical models, we see that except for the fact that the estimated efficiency scores are different, we obtain very different results regarding the efficiency rankings of the different manufacturing sectors. In particular, in Table 1.11 of Appendix B, we present the technical efficiency scores and the corresponding rankings for each manufacturing sector. Table 1.11, depicts the fact that using static models where the partial adjustment structure is ignored, will fairly distort the efficiency rankings of each sector.

1.6.5 Total Factor Productivity Analysis

Last, in this subsection we present estimates of the technological progress, efficiency change and the total factor productivity growth from the DRC-SFM. The technological progress is defined as the exogenously induced shifts of the production frontier captured by the parameter estimated of the time effects $\boldsymbol{f}_t$, the efficiency change between two period is defined as $-(u_{it} - u_{it-1})$ and the total factor productivity can be estimated as the sum of the two components. In Table 1.10, we present the estimated TFP growth and it's sub-components. In addition, we present the TFP and the Efficiency Index in order to illustrate the productivity and efficiency intertemporal behaviour.

Table 1.10: Technological Progress, Efficiency Change and TFP Growth

	Tech. Progress	Eff. Change	TFP Growth	TFP Index	Eff. Index
Pre Reform Period					
1987/88				100	100
1988/89	-1.59%	0.87%	-0.72%	99.28	100.87
1989/90	-0.74%	-3.20%	-3.95%	95.35	97.64
1990/91	0.10%	-1.33%	-1.23%	94.18	96.34
Post Reform Period					
1991/92	0.95%	1.80%	2.75%	96.77	98.07
1992/93	1.79%	3.11%	4.90%	101.53	101.12
1993/94	2.64%	-4.04%	-1.39%	100.11	97.04
1994/95	3.49%	1.31%	4.80%	104.92	98.31
1995/96	4.33%	-3.65%	0.68%	105.64	94.72

Overall, we see that the liberation reforms introduced during the early 90s increased the total factor productivity of the private manufacturing sector. Specifically, the TFP Index illustrates that during the study period, the TFP increased by 5.64%. Additionally, we see that the main driver of productivity growth was the exogenous technological progress induced by market liberalization. On the other hand, we see that the technical efficiency during the post-reform period reduced significantly (on average by 6%), indicating that the sectors could not utilize the whole available technology despite the positive technological shock.

1.7 Conclusions

In this study, we proposed a Bayesian hierarchical panel stochastic frontier model which accounts for the production adjustment costs. In particular, we built a dynamic panel random coefficient stochastic frontier model where the heterogeneity across the production units is captured by utilizing the fact that production systems can be subject to heterogeneous technological capabilities. We illustrate that our model can adequately separate the time-invariant unobserved heterogeneity effects from the latent time-varying inefficiency effects. To estimate the model, we present Bayesian Markov Chain Monte Carlo (MCMC) techniques to effectively obtain the posterior densities of the parameters of interest. The proposed Gibbs Sampler can be very easily implemented by empirical researchers. Using artificial data, we illustrate that our model performs very well in small and moderate samples.

In an empirical application, we show that static models that ignore the dynamics of a production process can generate misleading estimates regarding the short-run technical efficiency estimates. In addition, we illustrate that the market reforms introduced in Egypt during the early 90s, created significant factor adjustment frictions, with an adjustment rate of approximately 86%. Moreover, we show that the reform introduced during that period increased total factor productivity, which was mainly driven by exogenous technological progress. These results are in line with the recent economic literature which investigates the role of factor adjustment frictions in the production process, indicating that our reduced form model provides economic meaningful inference.

In terms of future research, an interesting extension would be to allow the probability density of the inefficiency term to be a function of different environmental variables and analyze the firm-level and macroeconomic determinants of technical inefficiency. Moreover, the model can be extended to allow for time variation in the autoregressive parameter and hence study the evolution of adjustment frictions in time.

Appendix A

In this Appendix, we present the Bayesian posterior analysis of the BH-TRE-SFM. The TRE-SFM takes the following form:

$$y_{it} = \boldsymbol{x}_{it}'\boldsymbol{\beta} + a_i + v_{it} - u_{it} \quad , \quad i = 1, 2, \ldots, N \quad , \quad t = 1, 2, \ldots, T$$

$$a_i \sim \mathcal{N}(0, \sigma_a^2)$$

$$v_{it} \sim \mathcal{N}(0, \sigma_v^2)$$

$$u_{it} \sim \mathcal{N}^{+}(0, \sigma_u^2)$$

given that $\boldsymbol{x}_{it}$ contains the vector with ones. In matrix form the model can be written as:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + (\boldsymbol{a} \otimes \mathbf{1}_T) + \boldsymbol{v} - \boldsymbol{u}$$

where

$$\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} \quad \text{and} \quad \mathbf{1}_T = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

As above, for the prior specifications of the parameters we have:

$$\pi(\beta) \propto 1$$

$$\sigma_a^2 | q_0, q_1 \sim \mathcal{IG}(q_0, q_1)$$

$$\sigma_v^2 | a_0, a_1 \sim \mathcal{IG}(a_0, a_1)$$

$$\sigma_u^2 | \gamma_0, \gamma_1 \sim \mathcal{IG}(\gamma_0, \gamma_1)$$

From the above, the joint augmented posterior distribution will be:

$$\begin{aligned} p(\boldsymbol{\Theta}|\boldsymbol{Y}, \boldsymbol{X}) \propto\; & (\sigma_v^2)^{-\frac{NT}{2}} exp\left\{ -\frac{1}{2\sigma_v^2} (\boldsymbol{Y} - \boldsymbol{X\beta} - (\boldsymbol{a} \otimes \mathbf{1}_T) + \boldsymbol{u})'(\boldsymbol{Y} - \boldsymbol{X\beta} - (\boldsymbol{a} \otimes \mathbf{1}_T) + \boldsymbol{u}) \right\} \\ & \times (\sigma_a^2)^{-\frac{N}{2}} exp\left\{ -\frac{1}{2\sigma_a^2} \sum_{i=1}^{N} a_i^2 \right\} \\ & \times (\sigma_u^2)^{-\frac{NT}{2}} exp\left\{ -\frac{1}{2\sigma_u^2} \sum_{i=1}^{N} \boldsymbol{u}_i'\boldsymbol{u}_i \right\} \mathbf{1}(\boldsymbol{u}_i \geq 0) \\ & \times (\sigma_a^2)^{-q_0 - 1} exp\left\{ -\frac{q_1}{\sigma_a^2} \right\} \\ & \times (\sigma_v^2)^{-a_0 - 1} exp\left\{ -\frac{a_1}{\sigma_v^2} \right\} \\ & \times (\sigma_u^2)^{-\gamma_0 - 1} exp\left\{ -\frac{\gamma_1}{\sigma_u^2} \right\} \end{aligned}$$

where the first line reflects that $\boldsymbol{Y}|\boldsymbol{X}, \boldsymbol{a}, \boldsymbol{u} \sim \mathcal{N}(\boldsymbol{X\beta} + (\boldsymbol{a} \otimes \mathbf{1}_T) - \boldsymbol{u}, \sigma_v^2)$, the second line reflects the normality of $a_i \sim N(0, \sigma_a^2)$, the third line depicts the half-normal distribution of the inefficiency term and the remaining three lines correspond to the prior specification of the parameters of interest.

From the augmented posterior distribution we can derive the conditional distributions of the parameters of interest. In particular, we have:

- **Conditional distribution of $\boldsymbol{\beta}$:**

$$\boldsymbol{\beta}|\boldsymbol{a}, \sigma_v^2, \boldsymbol{u} \sim \boldsymbol{\mathcal{N}}(\hat{\boldsymbol{b}}, \sigma_v^2(\boldsymbol{X}'\boldsymbol{X})^{-1})$$

  where

$$\hat{\boldsymbol{b}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{X}'(\boldsymbol{Y} - (\boldsymbol{a} \otimes \mathbf{1}_T) + \boldsymbol{u}))$$

- **Conditional distribution of σ_v^2:**

$$\sigma_v^2|\boldsymbol{\beta}, \boldsymbol{a}, \sigma_a^2, \sigma_u^2, \boldsymbol{u}, a_0, a_1 \sim \mathcal{IG}\left(\frac{NT}{2} + a_0, \frac{(\boldsymbol{Y} - \boldsymbol{X\beta} - (\boldsymbol{a} \otimes \mathbf{1}_T) + \boldsymbol{u})'(\boldsymbol{Y} - \boldsymbol{X\beta} - (\boldsymbol{a} \otimes \mathbf{1}_T) + \boldsymbol{u})}{2} + a_1\right)$$

- **Conditional distribution of σ_u^2:**

$$\sigma_u^2|\boldsymbol{\beta}, \boldsymbol{a}, \sigma_v^2, \sigma_a^2, \boldsymbol{u}, \gamma_0, \gamma_1 \sim \mathcal{IG}\left(\frac{NT}{2} + \gamma_0, \frac{\sum_{i=1}^N \boldsymbol{u}_i'\boldsymbol{u}_i}{2} + \gamma_1\right)$$

- **Conditional distribution of σ_a^2:**

$$\sigma_a^2|\boldsymbol{\beta}, \boldsymbol{a}, \sigma_v^2, \sigma_u^2, \boldsymbol{u}, q_0, q_1 \sim \mathcal{IG}\left(\frac{N}{2} + q_0, \frac{\sum_{i=1}^N a_i^2}{2} + q_1\right)$$

- **Conditional distribution of u_{it}:**

$$u_{it} \sim \mathcal{N}^+\left(-\frac{\sigma_u^2(y_{it} - a_i - x_{it}'\boldsymbol{\beta})}{\sigma_u^2 + \sigma_v^2}, \frac{\sigma_v^2\sigma_u^2}{\sigma_u^2 + \sigma_v^2}\right)$$

- **Conditional distribution of a_i:**

$$a_i \sim \mathcal{N}\left(\frac{1}{T + \frac{\sigma_v^2}{\sigma_a^2}}(\mathbf{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{u}_i)'\mathbf{1}_T, \frac{\sigma_v^2\sigma_a^2}{T\sigma_a^2 + \sigma_v^2}\right)$$

Appendix B

In this appendix, we illustrate differences regarding the technical efficiency estimates between the static and the dynamic Bayesian hierarchical model. In Table 1.11, we report the average technical efficiencies and the corresponding rankings for each sector for the BH-TRE-SFM the BH-DRC-SFM.

Table 1.11: Technical Efficiency estimates and rankings of Manufacturing Sectors for the different models

		BH-TRE-SFM		BH-DRC-SFM	
	Industrial Activity	Average TE	Ranking	Average TE	Ranking
1	Food Manufacturing	83.9%	4	84.6%	5
2	Other Food Manufacturing	84.0%	3	85.5%	2
3	Beverage and liquor	80.3%	27	83.6%	12
4	Tobacco	82.6%	16	82.9%	15
5	Manufacture of textile	83.6%	10	84.0%	8
6	Manufacture of wearing apparels	83.6%	8	85.3%	3
7	Manufacture of leather products	82.7%	15	83.6%	11
8	Manufacture of footwear	83.8%	5	84.6%	6
9	Manufacture of wood products	81.0%	25	82.2%	21
10	Manufacture of furniture & fixture	81.0%	24	82.6%	18
11	Manufacture of paper products	76.9%	28	82.4%	20
12	Printing and publishing industries	80.8%	26	80.1%	27
13	Manufacture of industrial chemicals	82.9%	13	81.4%	24
14	Manufacture of other chemical products	84.5%	2	87.3%	1
15	Other petroleum and coal	82.2%	20	82.9%	16
16	Manufacture of rubber products	81.8%	22	82.7%	17
17	Manufacture of plastic products	83.7%	7	83.2%	13
18	Manufacture of pottery and china	83.8%	6	83.6%	10
19	Manufacture of glass and glass products	84.6%	1	84.0%	8
20	Manufacture of other non metallic products	83.4%	11	82.4%	19
21	Iron and steel basic industries	81.2%	23	80.3%	26
22	Non-ferrous basic industries	82.4%	27	83.9%	9
23	Manufacture of fabricated metal products	82.8%	14	80.7%	25
24	Manufacture of machinery except electrical	83.6%	9	82.2%	2
25	Manufacture of electrical machinery	83.3%	12	83.1%	14
26	Manufacture of transport equipment	82.1%	21	84.9%	4
27	Manufacture of professional equipment	82.3%	19	79.6%	28
28	Other manufacture industries	82.3%	18	81.5%	23

Chapter 2

Dynamic Panel Stochastic Frontier Models with Inefficiency Effects

Abstract

This chapter proposes a dynamic panel production frontier model where the production process can be subject to production adjustment effects. The idea behind dynamic production functions is to allow input production and inefficiency shocks to have intertemporal effects on the production process. We extend the literature by incorporating the time-invariant unobserved heterogeneity and the time-varying technical inefficiency. The inefficiency term is specified as a parametric function of exogenous variables that may explain the technical inefficiency scores. A Generalized Method of Moments (GMM) approach is proposed where endogeneity issues can be effectively addressed. The proposed specification performs very well in small and moderate samples.

2.1 Introduction

Measuring firm productivity and technical efficiency is an important task for firms to assess their production performance, understand their cost-saving potentials, and, of course, help them to improve competitiveness and survive within a competitive market. Stochastic Frontier Analysis consists of a standard econometric tool to achieve this goal. Stochastic frontier models were first introduced by Aigner et al. (1977) and Meeusen and van Den Broeck (1977) and allow the econometric analysis to consider potential technical inefficiencies within the production process. Under this framework, the total factor productivity growth is decomposed into two parts, the common technological process, which affects all the firms symmetrically, and the technical efficiency change, which incorporates the firms' efficiency improvements throughout the years. Since its inception, stochastic frontier analysis has been used in a wide range of empirical analysis such as in the neoclassical production theory, banking models, educational economics and energy economics, among many others.

The most recent advances in panel data stochastic frontier models have focused on (i) effectively separating the time-invariant heterogeneity effects across the production units from the inefficiency effects and (ii) dealing with endogeneity problems, which are prominent in efficiency and productivity studies. Regarding the incidental parameter problem, the stochastic frontier literature has focused on ways to tackle this issue by proposing different ways to separate the latent heterogeneity from the inefficiency effects (see Tsionas (2002), Greene (2005a,b), Wang and Ho (2010), Chen et al. (2014), Belotti and Ilardi (2018), Kutlu et al. (2019) etc.), where the unobserved heterogeneity is modelled through production unit-specific effects or slope heterogeneity. In addition, some recent works that handle endogeneity problems in stochastic frontier models can be found at Kutlu (2010), Griffiths and Hajargasht (2016), Amsler et al. (2016), Kutlu et al. (2019), Centorrino and Pérez-Urdiales (2021), and Tsionas and Kumbhakar (2022) among others.

However, the majority of the aforementioned panel data stochastic frontier models lie on the assumption that once production inputs are introduced in the process, they are able to contribute to the final output within their maximum capabilities. Nevertheless, in real production environments, the instantaneous adjustment of inputs is a restrictive assumption. In the economics production theory, many papers highlight the importance of input adjustment costs. In more simple words, they support that once the inputs are introduced in the production system, it is required some time to get adapted to their new production environments. For example, there are periods when firms are facing unexpected demand shifts, and they require some time for their production supply to cover the new production demands. In addition, another fact is that the introduction of new inputs in the process, such as new employees (labour) or new production machines (capital), requires some additional time for employees to learn how to use these new machines, the time needed to train the new labour force, etc. Some more recent studies that illustrate the importance of accounting for the input adjustment frictions in the economic modelling consist of Hall (2004) where the author examines the factor adjustment costs for the US sectors, Cooper and Haltiwanger (2006) where capital adjustment costs are investigated in a sample of plant-level data, Groth and Khan (2010) examine the impact of investment costs on the US economy, Bergeaud and Ray (2021) explore the effects of real estate friction on firms' dynamics, and Artuç et al. (2010), Artuc et al. (2022) illustrate the effects of trade shocks on labour adjustment frictions, among others.

It is clear that in any production process, the sluggish behaviour of factor inputs can directly affect the production process and, consequently, the final produced output. As a result, model and econometric specifications which ignore the aforementioned production dynamics can generate misleading inferences regarding the short-run input elasticities and hence, the economic performance estimation of the production units.

This chapter proposes a simple stochastic frontier model that incorporates fixed effects and various forms of endogeneity along with the intertemporal behaviour of the production process. In particular, we propose a general dynamic panel production frontier model by incorporating the time-invariant unobserved heterogeneity across firms and the time-varying technical inefficiency. Furthermore, our model specification allows factor inputs to be either uncorrelated or correlated with the production shocks. Moreover, we assume a fully parameterized technical efficiency specification as proposed by Paul and Shankar (2018, 2019), where exogenous variables explain the level of technical inefficiency. To estimate the model, we propose a non-linear Generalized Method of Moments approach where arbitrary endogenous effects can be effectively addressed. We assess the proposed method using extensive Monte Carlo experiments and illustrate that the model performs very well in finite and moderate samples.

The rest of the chapter is organized as follows. In Section 2.2, we present the theoretical model. In Section 2.3, we present the econometric model. In Section 2.4, we present the proposed Non-Linear GMM. In Section 2.5, we illustrate some Monte Carlo experiments to assess the performance of the proposed model in small and moderate samples. Last, in Section 2.6, we conclude.

Related Literature: The proposed model fits into several literature strands as we extend existing specifications and bring several distinct threads into the literature together. First, Nickell (1996) and Nickell et al. (1997) are some of the existing papers in the literature that utilize dynamic panel production functions to incorporate the short-run adjustment frictions and examine the firms' total factor productivity. In all these studies, the authors analyze the total factor productivity through control variables that affect productivity evolution and do not explicitly estimate the corresponding technical inefficiencies. Moreover, Ayed-Mouelhi and Goaied (2003) and Bhattacharyya (2012) utilize the partial adjustment mechanism to model the intertemporal production behaviour allowing the presence of time-invariant technical inefficiency. As in our model, they utilize a dynamic panel model where the autoregressive parameter of the lag output captures the magnitude of the sluggish adjustment. They propose a two-stage solution, wherein the first step, a Generalized Method of Moments (GMM) approach, is used to estimate the parameters of interest. In order to retrieve the corresponding efficiency scores, the authors use a Schmidt and Sickles (1984) approach where the technical inefficiency is obtained by relative comparison with the "fully efficient" firm. In our model, in contrast, we allow technical inefficiency to be time-varying, and we separate the unobserved heterogeneity effects from the inefficiency effects. This is quite important in practice since, as has been illustrated in the stochastic frontier literature, failing to control for the unobserved heterogeneity adequately will result in very distorted efficiency scores (see Greene (2005a,b), Wang and Ho (2010), Chen et al. (2014), Belotti and Ilardi (2018), Kutlu et al. (2019), among others.).

Furthermore, Jonuzaj and Tsionas (2023) propose a flexible dynamic panel stochastic frontier model where the autoregressive parameter and the input elasticities can differ across the firms. They propose a Bayesian framework for model estimation and illustrate that the model can separate the firm-specific heterogeneity from the time-invariant inefficiency effects.

Another strand of the literature deals with reduced-form dynamic models, where the technical inefficiency is specified as an autoregressive function of its past values. The motivation behind the autoregressive structure of technical inefficiency is that the input adjustment costs will cause sluggish adoption of new technological innovations. Therefore, the technical inefficiency evolution towards the long-run state will be more stagnant. Econometric specification in this direction can be seen at Ahn et al. (2000), Tsionas (2006), Emvalomatis (2012a), Amsler et al. (2014) and Lai and Kumbhakar (2020).

Last, Tsionas et al. (2020) and Tsionas et al. (2022) provide some new structural models where the production process can be subject to adjustment costs. Although structural models can provide comprehensive economic results, in this paper, we focus on providing a flexible reduced-form model that empirical scholars can easily implement.

2.2 The Model

2.2.1 Theoretical Framework

Let's consider a production process where N firms are producing over a time period T. Each firm is using a set of inputs $X_{it} \in \mathbb{R}_+^p$ in order to produce a single output $Y_{it}^* \in \mathbb{R}_+$. In addition, the firm is subject to a Hicks-neutral total factor productivity level Ω_{it}. Output Y_{it}^* can be treated as a latent variable, not observed by the firms, and can be seen as the maximum feasible output or the targeted output that firms "wish" to succeed, conditional on the level of inputs and the productivity level. The production process can be described as:

$$Y_{it}^* = f(X_{it}; \beta)\Omega_{it} \tag{2.1}$$

where $f(.)$ can be any production function such as the Cobb-Douglas and β are the corresponding technological parameters. Following the stochastic frontier literature, we can decompose total factor productivity into three components; (i) a time-invariant firm-specific component that captures the production heterogeneity between firms, (ii) the common technological component that affects symmetrically all firms and (iii) a time-varying firm-specific component $\exp(-u_{it})$ which reflects the technological utilization or the technical efficiency of each firm. The technical inefficiency $\exp(-u_{it})$ lies on the interval $(0, 1]$ and reflects how efficient a firm is relative to the maximum potential production. Moreover, the inefficiency term u_{it}, should be non-negative and reflects the deviation of production from the maximum feasible output. In particular, we have:

$$\Omega_{it} = \exp(\eta_i + h_t - u_{it}) = \exp(\eta_i) \times \exp(h_t) \times \exp(-u_{it}) \tag{2.2}$$

Nevertheless, as mentioned above, firms can be subject to adjustment costs and are not able to adjust their production instantaneously to their targeted output Y_{it}^*. Instead, they face a gradual adjustment towards the desired level. The following equation can describe the mechanism behind the output adjustment:

$$\frac{Y_{it}}{Y_{it-1}} = \left(\frac{Y_{it}^*}{Y_{it-1}}\right)^{\lambda} \quad , \quad 0 < \lambda \leq 1 \tag{2.3}$$

In the economics literature, this mechanism is known as the partial adjustment model, and the idea behind this is that the observed output Y_{it} differs from the desired level Y_{it}^* due to some inertia that may exist in the production process. Equation 2.3 illustrates that the output adjustment is subject to the parameter λ which takes values between $(0, 1]$, and determines the magnitude of the adjustment. Here, we should note that parameter λ can not take the zero value, since we assume there must be some level of adjustment. Also, we assume that firms are identical to each other, they face the same production restrictions, and the level of adjustment is common across the firms. Hence, the parameter λ, can be seen as an adjustment speed of the production output to the desired one. For instance, in the case where $\lambda = 1$, firms adjust their outputs instantly, and apparently, the realized production output will be exactly the targeted one, viz. $Y_{it} = Y_{it}^*$. On the other hand, when $\lambda \to 0$, the speed of adjustment is relatively low, firms are not able to adjust their production supply to the demanded output, and the output gap increases.

From equation 2.3, by taking the natural logarithm, we end up with:

$$y_{it} - y_{it-1} = \lambda(y_{it}^* - y_{it-1}) \quad , \quad 0 < \lambda \leq 1 \tag{2.4}$$

where y_{it} is the natural logarithm of the observed output of firm i at time t, y_{it-1} is the natural logarithm of the output of firm i at the lagged period $t-1$, y_{it}^* is the targeted or the desired output that firms want to achieve and λ is the speed of adjustment.

Rewriting equation 2.4 and combining with equations 2.1 and 3.2, we have:

$$y_{it} = (1-\lambda)y_{it-1} + \lambda(\boldsymbol{x}_{it}'\boldsymbol{\beta} + \eta_i + \eta_t - u_{it}) + v_{it} \tag{2.5}$$

where v_{it} is usual two-sided error term. It is clear, that the presence of partial adjustment effects delivers a dynamic panel stochastic frontier model. In particular, equation 2.5 can be written more compactly as:

$$\begin{aligned} y_{it} &= \rho y_{it-1} + \boldsymbol{x}_{it}'\boldsymbol{\beta}^{sr} + \varepsilon_{it} \\ \varepsilon_{it} &= a_i + a_t + v_{it} - u_{it}^+ \end{aligned} \tag{2.6}$$

where $\rho = (1-\lambda)$, $\boldsymbol{\beta}^{sr} = \boldsymbol{\beta}\lambda$, $a_i = \lambda\eta_i$, $a_t = \lambda\eta_t$ and $u_{it}^+ = \lambda u_{it}$. Thus, the dynamic structure introduced in equations 2.5 and 2.6 allows input and inefficiency shocks to have an intertemporal impact on the current production process. As discussed above, this specification enables to model of the implicit dynamic behaviour of any production process. In particular, the proposed dynamic specification distinguishes between short-run and long-run effects of input and inefficiency shocks. Specifically, for the short and long-run input elasticities we have:

$$\text{SR:} \quad \frac{\partial y_{it}}{\partial \boldsymbol{x_{it}}} = \boldsymbol{\beta}^{sr} \qquad \text{LR:} \quad \frac{\partial y_{it}}{\partial \boldsymbol{x_{it}}} = \frac{\boldsymbol{\beta}^{sr}}{1-\rho} \tag{2.7}$$

and for inefficiency effects we have:

$$\text{SR:} \quad \frac{\partial y_{it}}{\partial u_{it}} = -1 \qquad \text{LR:} \quad \frac{\partial y_{it}}{\partial u_{it}} = -\frac{1}{1-\rho} \tag{2.8}$$

2.2.2 The Econometric Model

To generalize the above, the proposed Dynamic Panel Stochastic Frontier model (hereafter DPSF model) can be written in the following reduced form:

$$\begin{gathered} y_{it} = \rho y_{it-1} + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \varepsilon_{it} \ , \ 0 \leq \rho < 1 \\ \varepsilon_{it} = \eta_i + v_{it} - u^+_{it} \\ i = 1, 2, \ldots, N \ , \ t = 2, 3, \ldots, T \end{gathered} \tag{2.9}$$

The composite error term ε_{it} consists of η_i which captures the time-invariant unobserved heterogeneity across firms, $v_{it} \sim N(0, \sigma_v^2)$ is the two-sided error term which captures production shocks and u^+_{it} is the non-negative term which determines the technical inefficiency of firm i at time t. A common choice for technical inefficiency is to assume a distribution with non-negative support, such as the half-norm, the truncated normal, exponential, or gamma distribution. Moreover, we can allow inefficiency effects to be affected by different exogenous variables that may explain the inefficiency level as in Wang and Schmidt (2002), Wang and Ho (2010), among many others.

In this study, we follow Paul and Shankar (2018, 2019) where the technical efficiency term is specified as a parametric function of different exogenous variables that may explain the efficiency level. Specifically, given the fact that technical efficiency should lie on the $[0, 1]$ interval, the authors argue the use of cumulative distributions in order to ensure this property. In particular, we have:

$$TE_{it} = e^{-u^+_{it}} = H(\boldsymbol{z}'_{it}\boldsymbol{\gamma}) \Rightarrow -u^+_{it} = ln \ H(\boldsymbol{z}'_{it}\boldsymbol{\gamma}) \tag{2.10}$$

where $H(.)$ could be any cumulative distribution function, $\boldsymbol{z}_{it}$ is the vector of these exogenous variables and $\boldsymbol{\gamma}$ is the corresponding parameter vector. A common choice for the CDF could be the standard normal CDF or the Logistic function. Similar specifications have been utilized in many recent econometric and empirical studies, such as in Tsionas and Mamatzakis (2019) and Kumbhakar and Tsionas (2020).

Thus, combining the model in 2.9 with the specification in 2.10, we end up with the final form of the model:

$$y_{it} = \rho y_{it-1} + x'_{it}\beta + \eta_i + \eta_t + ln \ H(z'_{it}\gamma) + v_{it} \ , \ 0 \leq \rho < 1 \tag{2.11}$$

The model in equation 2.11 can be seen as a dynamic panel data model with non-linear inefficiency effects. In addition, specification 2.11 can be seen as a generalization of previous panel stochastic frontier models. In particular, under the assumption of instantaneous adjustment, viz. $\rho = 0$, the model collapse to the model proposed by Paul and Shankar (2018, 2019). The authors proposed a simple non-linear least squares approach and Tsionas and Mamatzakis (2019) argue for a GMM approach that can handle different endogeneity issues.

Moreover, under the full efficiency assumption, viz. $H(\boldsymbol{z}'_{it}\boldsymbol{\gamma}) = 1$, the model returns to a linear dynamic panel model. A number of methods have been proposed to estimate dynamic panel data, with the most prevailing in the empirical work being GMM methods. The main idea is that the presence of fixed effects will make the standard OLS and FE estimators inappropriate since the estimates of the autoregressive parameter will be biased in finite samples (see Nickell (1981)), and therefore we will have inaccurate estimates regarding the adjustment speed. For this reason, a standard approach is to eliminate the fixed effects using a first difference transformation, and then different instrumental variables can be used to consistently identify the parameter of interest. Estimation and inference in dynamic panel data models have been at the heart of modern econometrics and typically involve GMM methods. Relative literature consists of Holtz-Eakin et al. (1988), Anderson and Hsiao (1981, 1982), Arellano and Bond (1991) and Blundell and Bond (1998), where the authors proposed different GMM and SYS-GMM techniques. Recently, Cave et al. (2022) illustrated an extensive Monte Carlo simulation study, where the statistical performance of different dynamic panel estimators is evaluated.

Overall, consistent estimation of the non-linear dynamic panel stochastic frontier model in equation 2.11 will incorporate all the above. In the next section, we illustrate the proposed Non-Linear GMM approach.

2.3 The Non-Linear GMM approach

In this section, we illustrate the proposed Non-Linear GMM approach. For simplicity, we ignore the time effects from the model specification. Hence we have:

$$y_{it} = \eta_i + \rho y_{it-1} + \boldsymbol{x}'_{it}\boldsymbol{\beta} + ln\ H(\boldsymbol{z}'_{it}\boldsymbol{\gamma}) + v_{it} \tag{2.12}$$

where η_i are the usual fixed-effects, $\boldsymbol{x}'_{it} = [x_{1it}, x_{2it}, \ldots, x_{Pit}]$ is a $1 \times P$ vector of inputs, $\boldsymbol{z}'_{it} = [z_{1it}, z_{2it}, \ldots, z_{Kit}]$ is the $1 \times K$ vector of exogenous variables that affect the level of technical inefficiencies and v_{it} are uncorrelated production shocks outside firm's control.

To begin with, we assume for simplicity that the inputs are considered to be strictly exogenous, viz. $Cov(\boldsymbol{x}_{it}; v_{is}) = 0$ for all $s = 1, 2, \ldots, T$. In addition, we assume that the environmental vector $\boldsymbol{z}_{it}$ is strictly exogenous to production shocks, and hence $Cov(z_{it}; v_{is}) = 0$ for all $s = 1, 2, \ldots, T$. As discussed above, since the nuisance parameters η_i cause bias problems, from the first difference transformation, we end up with:

$$\Delta y_{it} = \rho \Delta y_{it-1} + \Delta \boldsymbol{x}'_{it}\boldsymbol{\beta} + m(\boldsymbol{z}'_{it}, \boldsymbol{z}'_{it-1}, \boldsymbol{\gamma}) + \Delta v_{it} \tag{2.13}$$

where

$$m(\boldsymbol{z}'_{it}, \boldsymbol{z}'_{it-1}, \boldsymbol{\gamma}) = \left[ln\ H(\boldsymbol{z}'_{it}\boldsymbol{\gamma}) - ln\ H(\boldsymbol{z}'_{it-1}\boldsymbol{\gamma})\right] \tag{2.14}$$

is the non-linear function of the strictly exogenous variables $\boldsymbol{z}'_{it}$ and $\boldsymbol{z}'_{it-1}$ which determines the first difference of technical inefficiencies. In order to identify the parameters in model 2.13 using the GMM approach, we need to define the moment conditions, as:

$$\mathbb{E}[\Delta v_{it}(\boldsymbol{\theta}) \otimes w_{it}] = 0 \tag{2.15}$$

where $\Delta v_{it}(\boldsymbol{\theta}) = \Delta y_{it} - \rho \Delta y_{it-1} - \Delta \boldsymbol{x}'_{it}\boldsymbol{\beta} - m(\boldsymbol{z}'_{it}, \boldsymbol{z}'_{it-1}, \boldsymbol{\gamma})$ and w_{it} is the vector of orthogonality conditions used to identify the parameters $\boldsymbol{\theta} = [\rho, \boldsymbol{\beta}, \boldsymbol{\gamma}]'$. As mentioned above, following Anderson and Hsiao (1981, 1982) and Arellano and Bond (1991) in order to identify the autoregressive parameter ρ, as instrumental variables we can use the level lagged values y_{it-s} with $s = 2, 3, \ldots, t-1$.

For identifying the input elasticities vector $\boldsymbol{\beta}$, inputs at first difference $\Delta \boldsymbol{x}_{it}$ can be considered as valid instruments since we assumed that they are strictly exogenous to the process and $\mathbb{E}(\Delta \boldsymbol{x}_{it} \Delta v_{it}) = 0$. Similarly, as long as $\boldsymbol{z}_{it}$ and $\boldsymbol{z}_{it-1}$ are assumed strictly exogenous, they can also be used to identify the corresponding parameter vector $\boldsymbol{\gamma}$. Thus, $w_{it} = [1, y_{it-s}, \Delta \boldsymbol{x}_{it}, \boldsymbol{z}_{it}, \boldsymbol{z}_{it-1}]'$ will be a $L \times 1$ vector of the instrumental variables.

Consequently, the moment conditions $g(\mathcal{Y}_{it}, \theta)$ will be given by:

$$g(\mathcal{Y}_{it}, \boldsymbol{\theta}) = \mathbb{E}\left[(\Delta y_{it} - \rho \Delta y_{it-1} - \Delta \boldsymbol{x}'_{it}\boldsymbol{\beta} - m(\boldsymbol{z}'_{it}, \boldsymbol{z}'_{it-1}, \boldsymbol{\gamma})) \otimes \begin{pmatrix} 1 \\ y_{it-s} \\ \Delta \boldsymbol{x}_{it} \\ \boldsymbol{z}_{it} \\ \boldsymbol{z}_{it-1} \end{pmatrix}\right] = 0 \tag{2.16}$$

where $\mathcal{Y}_{it} = [\Delta y_{it}, \Delta y_{it-1}, \Delta \boldsymbol{x}'_{it}, \boldsymbol{z}'_{it}, \boldsymbol{z}'_{it-1}]$ is the vector of the observed variables, $\theta \in \mathbb{R}^M$ is the vector of the unknown parameters ρ, β and γ and $s = 2, 3, \ldots, t-1$ determines the number of lagged values.

Here, we should highlight that the assumption of strictly exogenous production input is very restrictive. Overall, it is known that production and macroeconomic shocks can have an intertemporal effect on the firm's input decision. For this reason, we relax the strictly exogeneity assumption, and we allow v_{it} to have a dynamic impact on the input vector $\boldsymbol{x}_{it}$, viz. $Cov(\boldsymbol{x}_{it-s}; v_{it}) = 0$ for all $s \geq 0$. Under this assumption, the production inputs are considered to be predetermined, and therefore, the input choice will be affected by past production and macroeconomic shocks. Under this condition, the vector $\Delta\boldsymbol{x}_{it}$ can not be considered a valid instrument, since:

$$\begin{aligned} Cov(\Delta\boldsymbol{x}_{it};\Delta v_{it}) &= \mathbb{E}(\Delta\boldsymbol{x}_{it}\Delta v_{it}) \\ &= \mathbb{E}(x_{it}v_{it}) - \mathbb{E}(x_{it}v_{it-1}) - \mathbb{E}(x_{it-1}v_{it}) + \mathbb{E}(x_{it-1}v_{it-1}) \\ &= -\mathbb{E}(x_{it}v_{it-1}) \neq 0 \end{aligned}$$

To solve this correlation problem, we need to instrument $\Delta\boldsymbol{x}_{it}$. A valid instrument is the vector $\boldsymbol{x}_{it-1}$ since it is relevant with $\Delta\boldsymbol{x}_{it}$ and orthogonal to Δv_{it}, meaning $Cov(\boldsymbol{x}_{it-1};\Delta\boldsymbol{x}_{it}) \neq 0$ and $Cov(\boldsymbol{x}_{it-1};\Delta v_{it}) = 0$. Thus, under predetermined regressors, the moment conditions $g(\mathcal{Y}_{it},\theta)$ will be given by:

$$g(\mathcal{Y}_{it},\boldsymbol{\theta}) = \mathbb{E}\left[(\Delta y_{it} - \rho\Delta y_{it-1} - \Delta\boldsymbol{x}'_{it}\boldsymbol{\beta} - m(\boldsymbol{z}'_{it},\boldsymbol{z}'_{it-1},\boldsymbol{\gamma})) \otimes \begin{pmatrix} 1 \\ y_{it-s} \\ \boldsymbol{x}_{it-1} \\ \boldsymbol{z}_{it} \\ \boldsymbol{z}_{it-1} \end{pmatrix}\right] = 0 \tag{2.17}$$

Moreover, for any $l \geq 1$ we have $Cov(\boldsymbol{x}_{it-l};\Delta v_{it}) = 0$ and hence, one can increase the corresponding orthogonality conditions. However, when the number of periods is large, we must consider the many weak instrument problems.

Given the moment conditions in 2.16 and 2.17, in order to estimate the GMM parameters, the minimization criterion will be:

$$Q(\boldsymbol{\theta}) = arg\ min\ \left[\frac{1}{N(T-2)}\sum_{i=1}^{N}\sum_{t=1}^{T-2} g(\mathcal{Y}_{it},\boldsymbol{\theta})\right]' \mathcal{W} \left[\frac{1}{N(T-2)}\sum_{i=1}^{N}\sum_{t=1}^{T-2} g(\mathcal{Y}_{it},\boldsymbol{\theta})\right] \tag{2.18}$$

where $\mathcal{W}$ is a $(L \times L)$ positive definite weighting matrix which weights the moment conditions. For the first step, one can use the identity matrix $(W = I)$ by giving a-priori the same weights to all moment conditions. This is known as the 1-step GMM (D-GMM-1) and we can obtain consistent estimates and asymptotically normal. The variance-covariance matrix of the D-GMM-1 estimator will have the form:

$$\Sigma = (G'\mathcal{W}G)^{-1}G'\mathcal{W}\Omega\mathcal{W}G(G'\mathcal{W}G)^{-1} \tag{2.19}$$

where

$$G = \frac{1}{N(T-2)}\sum_{i=1}^{N}\sum_{t=1}^{T-2}\frac{\partial g(\mathcal{Y}_{it},\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\ ,\ \Omega = \frac{1}{N(T-2)}\sum_{i=1}^{N}\sum_{t=1}^{T-2} g(\mathcal{Y}_{it},\boldsymbol{\theta})g(\mathcal{Y}_{it},\boldsymbol{\theta})' \tag{2.20}$$

For the second step, it has be shown that by choosing the optimal weighting matrix, as $\hat{\mathcal{W}} = \hat{\Omega}^{-1}$ and plugged in the minimization criterion 2.18, we can obtain the most efficient estimator. This method is called 2-step GMM (D-GMM-2) and the estimator is still consistent and asymptotically normal, with variance-covariance matrix:

$$\Sigma = (G'\hat{\mathcal{W}}G)^{-1} \tag{2.21}$$

The same procedure as the 2-step GMM, can be performed iteratively, where the weighting matrix $\hat{\mathcal{W}}$ is recalculated several times until the estimator converges to a specific value. This method is called Iterated-GMM (D-GMM-Iter), and the minimization criterion, will be:

$$Q(\boldsymbol{\theta}_{b+1}) = arg\ min\ \left[\frac{1}{N(T-2)}\sum_{i=1}^{N}\sum_{t=1}^{T-2} g(\mathcal{Y}_{it}, \hat{\boldsymbol{\theta}}_b)\right]' \mathcal{W}(\hat{\boldsymbol{\theta}}_b)\ \left[\frac{1}{N(T-2)}\sum_{i=1}^{N}\sum_{t=1}^{T-2} g(\mathcal{Y}_{it}, \hat{\boldsymbol{\theta}}_b)\right] \tag{2.22}$$

where $b = 2, 3, \ldots, B$ is the number of iterations performed until convergence succeeds. However, the Iterated-GMM (D-GMM-Iter) does not provide any asymptotic improvement. Nevertheless, many Monte-Carlo experiments suggest this approach performs better in finite samples.

Last, an alternative approach is the Continuously-Updated GMM (D-GMM-CUE) estimator, which allows the weighting matrix to be a function of the parameter vector $\boldsymbol{\theta} = [\rho, \boldsymbol{\beta}, \boldsymbol{\gamma}]'$. The minimization criterion is:

$$Q(\boldsymbol{\theta}) = arg\ min\ \left[\frac{1}{N(T-2)}\sum_{i=1}^{N}\sum_{t=1}^{T-2} g(\mathcal{Y}_{it}, \boldsymbol{\theta})\right]' \mathcal{W}(\boldsymbol{\theta})\ \left[\frac{1}{N(T-2)}\sum_{i=1}^{N}\sum_{t=1}^{T-2} g(\mathcal{Y}_{it}, \boldsymbol{\theta})\right] \tag{2.23}$$

where both the parameter vector $\boldsymbol{\theta}$ and the optimal weighting matrix are $\mathcal{W}(\boldsymbol{\theta})$ are estimated simultaneously.

By minimizing the GMM criterion using all the aforementioned methods, we can obtain the estimated parameters $\hat{\boldsymbol{\theta}} = [\hat{\rho}, \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\gamma}}]'$. As a result, the estimation of the technical efficiency scores is straightforward, since $\hat{TE}_{it} = H(\boldsymbol{z}'_{it}\hat{\gamma})$.

2.4 Monte Carlo Replications

In this section, we investigate the performance of the proposed Non-Linear GMM method in finite samples. In particular, we assess the model under the two aforementioned scenarios: (i) the firm's inputs are strictly exogenous, viz. $Cov(x_{it}; v_{is}) = 0$ for all $s = 1, 2, \ldots, T$, which is a common assumption in the stochastic frontier literature; and (ii) the input decision is predetermined, viz. $Cov(x_{it-s}; v_{it}) = 0$ for all $s \geq 0$.

More specifically, the first scenario implies that production shocks do not affect the input decision of the firm. In other words, the firm is choosing its inputs according to an internal decision rule which is based on the firm's business plan. In the second scenario, we allow past production shocks to affect the current input decision of the firm. For instance, the impact of different shocks could take some time to affect the firm's decision for labour or investments.

We estimate the model for different number of firms $N = \{100, 200\}$ and different time periods $T = \{5, 10\}$. In addition, we estimate the model for different values of the autoregressive parameter $\rho = \{0.2, 0.5, 0.8\}$. In order to guarantee a safe inference of the results, we use 500 Monte Carlo iterations. We assess the non-linear GMM approach based on the estimated frontier parameters and the estimated technical efficiencies.

For the purpose of this experiment, we estimate a simple AR(1) dynamic panel production frontier model according to the following Data Generated Process (DGP):

$$\begin{gathered} y_{it} = \rho y_{it-1} + \beta x_{it} + \eta_i + v_{it} - h(z_{it}; \gamma)^+ \\ x_{it} = (1 - \rho_x)(\mu_x + \tau\eta_i) + \rho_x x_{it-1} + \xi_{it} + k v_{it-1} \\ h(z_{it}; \gamma)^+ = -ln\ \Phi(\gamma_1 z_{it}) \\ \eta_i \sim N(0, \sigma_\eta^2)\ ,\ v_{it} \sim N(0, \sigma_v^2)\ ,\ \xi_{it} \sim N(0, \sigma_\xi^2) \end{gathered} \tag{2.24}$$

where η_i captures the time-invariant unobserved heterogeneity across firms, y_{it} is the natural logarithm of the observed output of firm i at t, y_{it-1} is the natural logarithm of firm i at time $t-1$, x_{it} is the natural logarithm of input used by firm i at time t, $h(z_{it}; \gamma)^+$ is the non-negative function which represents the inefficiency of firm i at time t, z_{it} is an exogenous variable which affects firm's technical efficiency scores and v_{it} is the two-sided error term representing the production and general macroeconomic shocks outside firms control.

The input used in the process x_{it} is generated as an AR(1) function of the lagged input x_{it-1}, in order to allow deviations from the long-run mean. The long-run state is given by the term $\mu_x + \tau\eta_i$. Hence, we allow heterogeneity effects to be correlated with x_{it}. In addition, we allow lag production and business cycle shock v_{it-1} to affect input decisions, as well. Specifically, v_{it-1} is augmented with a constant parameter k, which determines the magnitude of this effect.

For instance, when $k = 0$, production shocks do not affect input choice, and hence the input x_{it} is strictly exogenous to the process. On the other hand, when $k \neq 0$, input choice is affected by the past production shocks and therefore, becomes predermined to the production process. We pick $\mu_x = 1$, $\tau = 0.25$ and $\rho_x = 0.7$ which indicates a persistent process. Moreover, we add a stochastic term ξ_{it} in order to capture unobserved factors with $\sigma_\xi = 0.1$.

For the technical efficiency, we assume a single exogenous factor z_{it} which is generated as $z_{it} = t/max(T) + w_{it}$, with $w_{it} \sim N(0, 0.1^2)$. Here, we want to generate a non-negative variable that is increasing in time, because we want to represent different factors, such as managerial skills or experience, which overall are represented by non-negative variables. In addition, we assume $\gamma_1 = 1$. A positive γ_1 parameter indicates that the variable z_{it} positively affects the technical efficiency scores. We choose these variables in order to generate technical efficiencies between 70% and 95%. This generated range of technical efficiency scores is common in many empirical findings.

For the time-invariant unobserved heterogeneity, we draw from a $N(0, \sigma_\eta^2)$ with $\sigma_\eta = 1$. To generate the two-sided error term we draw from a $N(0, \sigma_v^2)$ with $\sigma_v = 0.1$. In addition, for the initial values we assume $y_{i0} = \frac{\beta}{1-\rho} x_{i0}$ and $x_{i0} = \mu_x + \tau\eta_i$.

In order to ensure the existence of the above effects in our generated data, for each firm i we generate $m+T$ time periods. Then, for each firm, the first time $m = 20$ periods are dropped from the analysis. Last, the proposed model is estimated using the 1-step GMM (D-GMM-1), the 2-step GMM (D-GMM-2), the Iterated GMM (D-GMM-Iter) and the Continuously updating GMM (D-GMM-CUE).

2.4.1 Model with strictly exogenous inputs ($k = 0$)

As discussed above, for the input x_{it} to be strictly exogenous to the process, we need to impose $k = 0$. As a result, the input is unaffected by past production shocks v_{it-1}. In order to estimate the model, we use the following moment conditions:

$$\mathbb{E}\left[(\Delta y_{it} - \rho \Delta y_{it-1} - \beta \Delta x_{it} - m(z_{it}, z_{it-1}; \gamma)) \otimes \begin{pmatrix} 1 \\ y_{it-s} \\ \Delta x_{it} \\ z_{it} \\ z_{it-1} \\ z_{it}^2 \\ z_{it-1}^2 \end{pmatrix} \right] = 0 \tag{2.25}$$

Here, for simplicity we use $s = 2$, indicating that the only instrument used in order to identify the autoregressive parameter ρ is the lagged value y_{it-2}. Apparently, one can use more lagged values, such as $s = 2, 3, \ldots, T-1$. In addition, in order to help the estimation procedure, we adopt a common practise in GMM approach, by adding more moment conditions, such as the z_{it}^2 and z_{it-1}^2. Therefore, the moment conditions in 2.25 consist of a system of 7 non-linear equations with 4 unknown parameters.

In Table 2.1 and Table 2.2, we report the Bias and the Mean Squared Error (hereafter MSE) of the estimated parameters for $N = 100$ and $N = 200$, respectively. Each table consists of three panels (Panel A, Panel B and Panel C), where we present the simulation results for the different values of the autoregressive parameter ($\rho = 0.2$, $\rho = 0.5$ and $\rho = 0.8$). We generate parameter β as $\beta = 1 \times (1-\rho)$ such that the long-run effect equals 1.

It is clear, that the proposed non-linear GMM performs well in finite samples. In particular, we see that as we increase the time period T, the MSE of the parameters decreases. The same behaviour can be observed, when we keep the time constant and increase the number of firms N. In addition, we see that the Biases of the parameters are very low, as well.

Moreover, in Figure 2.1, we present some indicative results for technical efficiencies estimates using the results from Panel A presented in Tables 2.1 and 2.2. In particular, for each model, we estimate the technical efficiency as $\Phi(\hat{\gamma}_1 z_{it})$, where $\hat{\gamma}_1 = \frac{1}{S}\sum_{s=1}^{S} \hat{\gamma}_1^s$ and $\hat{\gamma}_1^s$ is the estimated parameter from each iteration $s = 1, 2, \ldots, 500$. Overall, all the different models generate technical efficiency estimates very close to the real ones. An interesting fact is that not only we are able to obtain accurate estimates regarding the average technical efficiency score, but also we can capture the whole shape of the efficiency density.

Last, the 2-step GMM (D-GMM-2) outperforms the 1-step GMM (D-GMM-1). Specifically, we obtain smaller Biases and MSE for all the different parameters, number of time periods T and number of firms N. In addition, the D-GMM-Iter and D-GMM-CUE seem to perform very well in finite samples, but they do not outperform the D-GMM-2. Overall, our Monte Carlo results indicate that all the proposed methods can be utilized effectively for empirical analysis.

2.4.2 Model with predetermined inputs ($k \neq 0$)

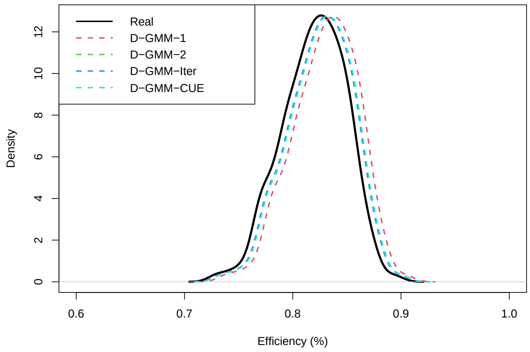
Following the above, to generate endogenous input, we need $k \neq 0$. For this simulation, we impose $k = 0.1$. Since $\mathbb{E}[x_{it} v_{it-1}] \neq 0$, in order to identify the parameter β, the lagged values x_{it-1} can be used as valid instruments. It is clear, one can use more lagged values, such as $x_{t-1}, x_{t-2}, \ldots, x_{i1}$. However, for illustration purposes, we choose to use only one lagged value, to keep our analysis as simple as possible. Hence, in this case, to estimate the model, we make use of the following moment conditions:

$$\mathbb{E}\left[(\Delta y_{it} - \rho\Delta y_{it-1} - \beta\Delta x_{it} - m(z_{it}, z_{it-1};\gamma)) \otimes \begin{pmatrix} 1 \\ y_{it-s} \\ x_{it-1} \\ z_{it} \\ z_{it-1} \\ z_{it}^2 \\ z_{it-1}^2 \end{pmatrix}\right] = 0 \tag{2.26}$$

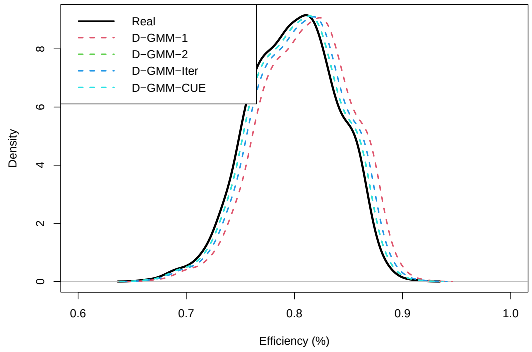
Again, to identify the autoregressive parameter ρ, for simplicity, we use $s = 2$. As we see in 2.26, except for the orthogonality condition for identifying β, all the remaining moment conditions are the same as in 2.25. As before, in Table 2.3 and Table 2.4, we report the Bias and the MSE of the estimated parameters and the estimated technical efficiencies for both $N = 100$ and $N = 200$.

Again, the proposed non-linear GMM performs very well, even in the presence of predetermined regressors. In particular, we see that the MSE of the estimated parameters is decreasing in firm size N and time period T. Moreover, in Figure 2.2, we present the kernel densities of the real efficiencies (black solid line) and the estimated efficiency scores using the corresponding methods. The reported densities indicate that all methods generate very accurate efficiency score estimates. Overall, we can conclude that our proposed GMM model can effectively address arbitrary endogeneity issues.

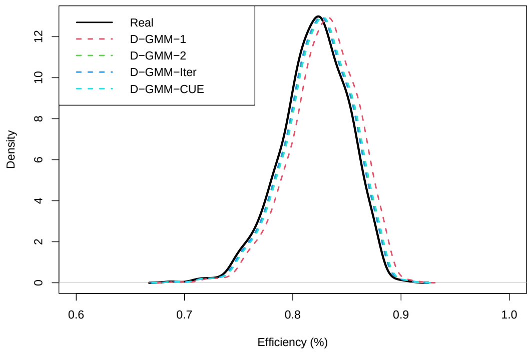
Last, the D-GMM-2 method performs better than the D-GMM-1 which aligns with the GMM theoretical properties. Moreover, the D-GMM-Iter and the D-GMM-CUE produce very small MSEs for all parameters of interest, which indicates that they can be good alternatives in empirical applications.



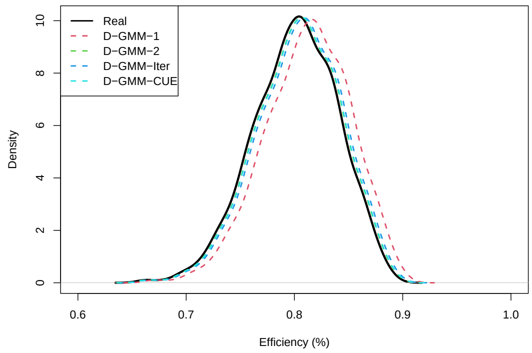
(a) $N = 100$, $T = 5$



(b) $N = 100$, $T = 10$

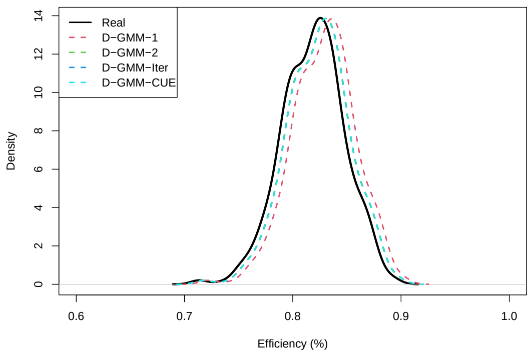


(c) $N = 200$, $T = 5$

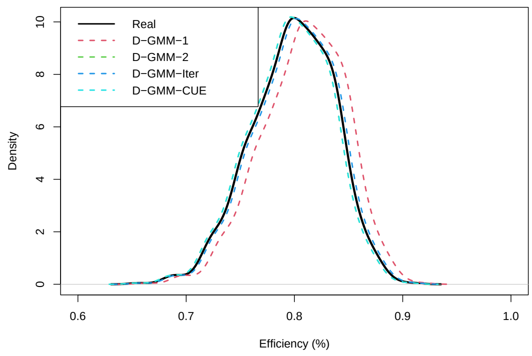


(d) $N = 200$, $T = 10$

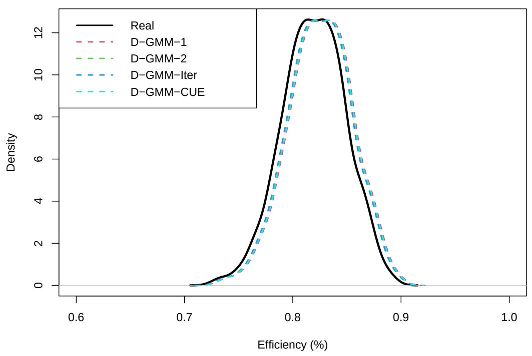
Figure 2.1: Kernel Densities of Estimated Technical Efficiencies for $\rho = 0.2$ and $k = 0$



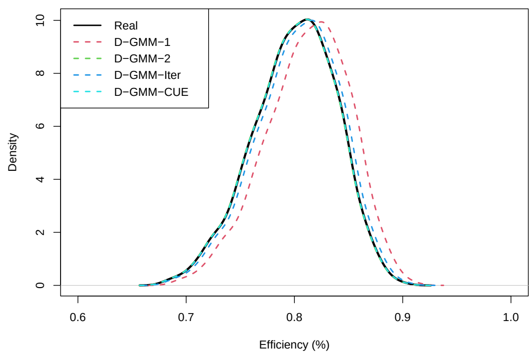
(a) $N = 100$, $T = 5$



(b) $N = 100$, $T = 10$



(c) $N = 200$, $T = 5$



(d) $N = 200$, $T = 10$

Figure 2.2: Kernel Densities of Estimated Technical Efficiencies for $\rho = 0.2$ and $k = 0.1$

2.5 Conclusions

In this chapter, we propose a dynamic panel production frontier model where the production process can be subject to partial adjustment effects. We extend the literature by allowing factor inputs and inefficiency shocks to have an intertemporal impact on production. Furthermore, our model is able to incorporate the time-invariant unobserved heterogeneity along with the time-varying technical inefficiency. The technical efficiency term is specified using cumulative distribution functions of exogenous variables that may explain the efficiency scores. A Generalized Method of Moments (GMM) approach can be used where endogeneity issues can be effectively addressed. We evaluate the performance of the proposed model under different scenarios, where inputs can be both exogenous and predetermined to the model. We show that in both cases, our model performs very well in finite samples.

In terms of future work, we plan to illustrate the usefulness of the proposed model in an empirical application where we will use firm-level data for the UK manufacturing and construction sector. The application will be interesting to examine the impact of certain firm-level and macroeconomic characteristics on the level of technical efficiency. In addition, part of the future research agenda is to extend the model and allow the autoregressive parameter to be a time-varying function of different variables. This will allow us to exploit the determinants of the sluggish adjustment of production.

Table 2.1: Monte Carlo Experiment for strictly exogenous input and $N = 100$

Panel A: $\rho = 0.2$, $\beta = 0.8$, $\gamma_1 = 1$							
		ρ		β		γ_1	
		Bias	MSE	Bias	MSE	Bias	MSE
$T = 5$	D-GMM-1	0.0026	0.0212	0.0030	0.0219	0.0462	0.0236
	D-GMM-2	-0.0355	0.0179	-0.0007	0.0061	0.0256	0.0274
	D-GMM-Iter	-0.0400	0.0184	-0.0010	0.0063	0.0222	0.0278
	D-GMM-CUE	0.0016	0.0207	0.0038	0.0066	0.0264	0.0273
$T = 10$	D-GMM-1	-0.0008	0.0091	-0.0012	0.0071	0.0532	0.0158
	D-GMM-2	-0.0156	0.0082	-0.0021	0.0024	0.0141	0.0252
	D-GMM-Iter	-0.0193	0.0086	-0.0025	0.0024	0.0294	0.0258
	D-GMM-CUE	-0.0072	0.0093	-0.0010	0.0025	0.0144	0.0255

Panel B: $\rho = 0.5$, $\beta = 0.5$, $\gamma_1 = 1$							
		ρ		β		γ_1	
		Bias	MSE	Bias	MSE	Bias	MSE
$T = 5$	D-GMM-1	0.0028	0.0173	0.0002	0.0257	0.0430	0.0243
	D-GMM-2	-0.0282	0.0160	-0.0037	0.0062	0.0282	0.0289
	D-GMM-Iter	-0.0313	0.0163	-0.0035	0.0063	0.0198	0.0291
	D-GMM-CUE	0.0046	0.0178	0.0005	0.0065	0.0316	0.0285
$T = 10$	D-GMM-1	0.0030	0.0050	0.0018	0.0101	0.0320	0.0158
	D-GMM-2	-0.0073	0.0044	-0.0016	0.0024	0.0008	0.0263
	D-GMM-Iter	-0.0083	0.0045	-0.0018	0.0024	0.0197	0.0265
	D-GMM-CUE	0.0022	0.0047	-0.0005	0.0025	0.0023	0.0264

Panel C: $\rho = 0.8$, $\beta = 0.2$, $\gamma_1 = 1$							
		ρ		β		γ_1	
		Bias	MSE	Bias	MSE	Bias	MSE
$T = 5$	D-GMM-1	-0.0025	0.0057	-0.0044	0.0328	0.0110	0.0242
	D-GMM-2	-0.0111	0.0047	-0.0013	0.0062	0.0013	0.0270
	D-GMM-Iter	-0.0115	0.0048	-0.0011	0.0064	-0.0159	0.0272
	D-GMM-CUE	0.0012	0.0049	-0.0003	0.0066	0.0085	0.0273
$T = 10$	D-GMM-1	0.0008	0.0001	0.0071	0.0253	0.0018	0.0200
	D-GMM-2	-0.0010	0.0007	-0.0004	0.0027	0.0079	0.0250
	D-GMM-Iter	-0.0009	0.0007	-0.0004	0.0027	0.0027	0.0252
	D-GMM-CUE	0.0017	0.0007	-0.0002	0.0028	0.0076	0.0252

Table 2.2: Monte Carlo Experiment for strictly exogenous input and $N = 200$

Panel A: $\rho = 0.2$, $\beta = 0.8$, $\gamma_1 = 1$							
		ρ		β		γ_1	
		Bias	MSE	Bias	MSE	Bias	MSE
$T = 5$	D-GMM-1	0.0070	0.0177	0.0023	0.0167	0.0379	0.0212
	D-GMM-2	-0.0195	0.0136	0.0013	0.0035	0.0136	0.0278
	D-GMM-Iter	-0.0225	0.0139	0.0009	0.0035	0.0105	0.0282
	D-GMM-CUE	0.0031	0.0158	0.0043	0.0037	0.0178	0.0277
$T = 10$	D-GMM-1	0.0063	0.0041	-0.0032	0.0039	0.0504	0.0129
	D-GMM-2	-0.0059	0.0036	-0.0036	0.0012	0.0095	0.0236
	D-GMM-Iter	-0.0073	0.0037	-0.0038	0.0012	0.0197	0.0239
	D-GMM-CUE	-0.0017	0.0038	-0.0031	0.0012	0.0116	0.0238

Panel B: $\rho = 0.5$, $\beta = 0.5$, $\gamma_1 = 1$							
		ρ		β		γ_1	
		Bias	MSE	Bias	MSE	Bias	MSE
$T = 5$	D-GMM-1	0.0105	0.0112	0.0034	0.0198	0.0311	0.0224
	D-GMM-2	-0.0152	0.0097	-0.0003	0.0032	0.0132	0.0263
	D-GMM-Iter	-0.0169	0.0099	-0.0005	0.0032	0.0063	0.0266
	D-GMM-CUE	0.0042	0.0101	0.0019	0.0033	0.0146	0.0263
$T = 10$	D-GMM-1	0.0062	0.0024	-0.0027	0.0054	0.0322	0.0121
	D-GMM-2	-0.0021	0.0022	-0.0008	0.0013	0.0117	0.0228
	D-GMM-Iter	-0.0026	0.0023	-0.0009	0.0013	0.0228	0.0229
	D-GMM-CUE	0.0020	0.0023	-0.0004	0.0013	0.0131	0.0229

Panel C: $\rho = 0.8$, $\beta = 0.2$, $\gamma_1 = 1$							
		ρ		β		γ_1	
		Bias	MSE	Bias	MSE	Bias	MSE
$T = 5$	D-GMM-1	-0.0038	0.0027	-0.0028	0.0305	0.0135	0.0237
	D-GMM-2	-0.0068	0.0022	0.0023	0.0031	0.0191	0.0255
	D-GMM-Iter	-0.0073	0.0022	0.0026	0.0032	-0.0110	0.0260
	D-GMM-CUE	-0.0010	0.0022	0.0030	0.0032	0.0218	0.0258
$T = 10$	D-GMM-1	-0.0002	0.0005	-0.0073	0.0168	0.0025	0.0152
	D-GMM-2	-0.0005	0.0004	-0.0004	0.0012	0.0081	0.0242
	D-GMM-Iter	-0.0004	0.0004	-0.0004	0.0012	0.0151	0.0244
	D-GMM-CUE	0.0007	0.0004	-0.0003	0.0012	0.0099	0.0245

Table 2.3: Monte Carlo Experiment for predetermined input and $N = 100$

Panel A: $\rho = 0.2$, $\beta = 0.8$, $\gamma_1 = 1$							
		ρ		β		γ_1	
		Bias	MSE	Bias	MSE	Bias	MSE
$T = 5$	D-GMM-1	-0.0007	0.0234	-0.0252	0.0297	0.0428	0.0222
	D-GMM-2	-0.0437	0.0133	-0.0413	0.0191	0.0192	0.0252
	D-GMM-Iter	-0.0474	0.0139	-0.0460	0.0199	0.0211	0.0260
	D-GMM-CUE	-0.0138	0.0148	-0.0113	0.0229	0.0207	0.0256
$T = 10$	D-GMM-1	0.0043	0.0090	0.0005	0.0218	0.0457	0.0144
	D-GMM-2	-0.0119	0.0061	-0.0106	0.0122	-0.0116	0.0248
	D-GMM-Iter	-0.0145	0.0065	-0.0134	0.0128	0.0104	0.0253
	D-GMM-CUE	-0.0007	0.0074	0.0034	0.0147	-0.0116	0.0250

Panel B: $\rho = 0.5$, $\beta = 0.5$, $\gamma_1 = 1$							
		ρ		β		γ_1	
		Bias	MSE	Bias	MSE	Bias	MSE
$T = 5$	D-GMM-1	-0.0081	0.0158	-0.0094	0.0329	0.0335	0.0242
	D-GMM-2	-0.0315	0.0130	-0.0261	0.0188	0.0192	0.0269
	D-GMM-Iter	-0.0336	0.0133	-0.0275	0.0192	0.0164	0.0274
	D-GMM-CUE	0.0024	0.0144	-0.0021	0.0206	0.0245	0.0270
$T = 10$	D-GMM-1	0.0005	0.0042	0.0047	0.0245	0.0380	0.0162
	D-GMM-2	-0.0103	0.0041	-0.0102	0.0090	-0.0011	0.0271
	D-GMM-Iter	-0.0115	0.0043	-0.0117	0.0093	0.0185	0.0274
	D-GMM-CUE	-0.0003	0.0045	-0.0012	0.0100	0.0011	0.0272

Panel C: $\rho = 0.8$, $\beta = 0.2$, $\gamma_1 = 1$							
		ρ		β		γ_1	
		Bias	MSE	Bias	MSE	Bias	MSE
$T = 5$	D-GMM-1	-0.0014	0.0045	-0.0024	0.0346	0.0208	0.0244
	D-GMM-2	-0.0105	0.0039	-0.0075	0.0132	0.0105	0.0268
	D-GMM-Iter	-0.0112	0.0040	-0.0087	0.0135	-0.0017	0.0268
	D-GMM-CUE	0.0016	0.0040	-0.0007	0.0145	0.0134	0.0269
$T = 10$	D-GMM-1	0.0017	0.0009	0.0056	0.0309	0.0038	0.0178
	D-GMM-2	-0.0003	0.0007	-0.0025	0.0055	0.0006	0.0261
	D-GMM-Iter	-0.0002	0.0007	-0.0025	0.0055	0.0038	0.0261
	D-GMM-CUE	0.0026	0.0007	-0.0002	0.0057	0.0010	0.0264

Table 2.4: Monte Carlo Experiment for predetermined input and $N = 200$

Panel A: $\rho = 0.2$, $\beta = 0.8$, $\gamma_1 = 1$							
		ρ		β		γ_1	
		Bias	MSE	Bias	MSE	Bias	MSE
$T = 5$	D-GMM-1	0.0100	0.0153	-0.0076	0.0290	0.0295	0.0200
	D-GMM-2	-0.0246	0.0094	-0.0261	0.0182	0.0241	0.0259
	D-GMM-Iter	-0.0272	0.0099	-0.0281	0.0188	0.0207	0.0263
	D-GMM-CUE	-0.0036	0.0107	-0.0014	0.0206	0.0263	0.0257
$T = 10$	D-GMM-1	0.0088	0.0048	0.0044	0.0185	0.0534	0.0133
	D-GMM-2	-0.0025	0.0035	-0.0038	0.0070	-0.0022	0.0237
	D-GMM-Iter	-0.0038	0.0037	-0.0053	0.0074	0.0164	0.0241
	D-GMM-CUE	0.0023	0.0041	0.0027	0.0081	-0.0003	0.0240

Panel B: $\rho = 0.5$, $\beta = 0.5$, $\gamma_1 = 1$							
		ρ		β		γ_1	
		Bias	MSE	Bias	MSE	Bias	MSE
$T = 5$	D-GMM-1	-0.0008	0.0109	-0.0148	0.0320	0.0400	0.0222
	D-GMM-2	-0.0302	0.0098	-0.0281	0.0153	0.0132	0.0262
	D-GMM-Iter	-0.0318	0.0101	-0.0288	0.0155	0.0107	0.0263
	D-GMM-CUE	-0.0106	0.0103	-0.0114	0.0164	0.0152	0.0261
$T = 10$	D-GMM-1	0.0024	0.0022	-0.0015	0.0193	0.0309	0.0129
	D-GMM-2	-0.0067	0.0024	-0.0037	0.0044	0.0137	0.0245
	D-GMM-Iter	-0.0073	0.0024	-0.0043	0.0045	0.0209	0.0246
	D-GMM-CUE	-0.0024	0.0025	0.0003	0.0046	0.0160	0.0247

Panel C: $\rho = 0.8$, $\beta = 0.2$, $\gamma_1 = 1$							
		ρ		β		γ_1	
		Bias	MSE	Bias	MSE	Bias	MSE
$T = 5$	D-GMM-1	-0.0018	0.0022	-0.0090	0.0342	0.0154	0.0214
	D-GMM-2	-0.0051	0.0019	0.0002	0.0096	0.0057	0.0260
	D-GMM-Iter	-0.0055	0.0019	0.0002	0.0097	-0.0096	0.0261
	D-GMM-CUE	0.0011	0.0019	0.0056	0.0101	0.0094	0.0258
$T = 10$	D-GMM-1	0.0023	0.0005	0.0035	0.0287	0.0032	0.0153
	D-GMM-2	0.0003	0.0003	0.0011	0.0029	-0.0025	0.0238
	D-GMM-Iter	0.0003	0.0003	0.0012	0.0029	-0.0013	0.0238
	D-GMM-CUE	0.0015	0.0003	0.0022	0.0030	-0.0024	0.0240

Chapter 3

A simple method for modelling the energy efficiency rebound effects with an application to energy demand frontiers

Abstract

This chapter proposes a new simple approach to model the macroeconomic energy efficiency rebound effects. The method follows recent developments in stochastic frontier models and assumes that energy efficiency improvements do not necessarily reduce energy demand proportionally. Instead, we allow country-specific rebound effects to mitigate or intensify the efficiency effects on the aggregate energy demand. The method incorporates a reduced-form stochastic frontier model with country-specific heteroscedasticity, and the method of moments approach for estimating the country-specific rebound effects. The estimation can be implemented relatively easily in any standard statistical package. Last, we illustrate the model in an empirical application, estimating the energy efficiency scores and the corresponding rebound effects for 20 OECD member countries from 1980 to 2018. Our results reveal that for most countries, we find modest to considerable partial rebound effects ranging from 28% to 92%. In addition, we show that for 2018 the average energy efficiency score is approximately 84%, indicating that there is potential for further energy savings.

3.1 Introduction

Improving energy efficiency is a key strategic objective for developed economies around the globe to reduce their aggregate energy use, reduce their GHG emissions, and enhance their energy supply security by reducing their energy dependency on third parties. However, despite all the energy efficiency actions and the legislation implemented throughout recent years, it is evident that in many economies, the aggregate energy consumption remains at the same levels or, in some instances, continues to increase steadily at the same rates.

To incorporate the underlying drivers of aggregate energy demand, the energy economics literature has been extended in various directions trying to shed light on the different mechanisms that drive aggregate energy use. One interesting mechanism that has attracted the attention of many energy economists and policymakers is the so-called economy-wide rebound effect. The mechanism implies that part of the initial energy efficiency gains will be offset, and consequently, the markets cannot utilize all the energy efficiency actions. The intuition behind the macroeconomic rebound effects is that improving energy efficiency in the markets will decrease the cost of energy-intensive products and services, and consequently, this will generate additional disposable income and productivity gains. Thus, this will stimulate the demand for other products and services and, as a result, the demand for energy use (Saunders (2000)). However, the magnitude of this effect may vary for different economies and is subject to different factors related to the economic structure of each economy, behavioural factors and social norms.

Thus, estimating and assessing the magnitude of the energy efficiency rebound effects is essential for energy policymakers and contributes to understanding and designing the necessary policy actions. Although there is growing literature on empirically investigating the economy-wide energy rebound effects, the empirical strategy seems ambiguous among energy economists and policymakers (Stern (2020), and Brockway et al. (2021) present some excellent literature reviews).

This study tries to shed additional light on the topic by proposing an alternative econometric identification strategy. In this study, we follow recent advances in the literature and introduce an energy input distance function where we allow the inefficiency term to affect aggregate energy consumption in a non-proportional way. Specifically, we present a flexible econometric specification which allows for country-specific rebound effects. Our model does not impose any restriction on the sign and the magnitude of the rebound effect. The proposed specification implies a reduced form stochastic frontier model with country-specific heteroscedastic inefficiency term. We suggest a simple estimation strategy that involves two stages to avoid computational complexity. First, we propose estimating the reduced-form model using standard Maximum Likelihood or Bayesian techniques. In particular, we utilize Bayesian techniques organized around MCMC inference and illustrate the corresponding conditional distributions. Next, we construct relevant moment conditions using information from the posterior estimates and the theoretical moment conditions arising from the inefficiency distribution, which incorporate the energy efficiency rebound effects. Finally, we estimate the country-specific rebound effects in the second stage using the Generalized Method of Moments (GMM) approach.

Last, we present an empirical application where we illustrate the practicality of our method in a panel of 20 OECD member countries. Our empirical findings reveal that for most of the OECD economies, there is evidence in favour of considerable partial rebound effects ranging from approximately 28% to 92%. This supports the argument that a vast part of initial energy efficiency gains is re-spent in supporting economic development. We find that Denmark is the only economy that exhibits zero rebound effects. On the other hand, we find that during the whole sample period, Germany and U.K. have been conservative regarding their aggregate energy use. Regarding energy efficiency estimates, our study reveals that the average energy efficiency for 2018 is approximately 86%, but the values of energy efficiencies are quite heterogeneous, ranging from 70% to 99%.

The rest of the chapter is organized as follows. In section 3.2, we briefly present the definition of the economy-wide rebound effects. In section 3.3, we present the proposed econometric strategy. In section 3.4, we present the empirical study. Last, in section 3.5, we conclude.

Related Literature This chapter stands in the energy economics literature, which provides econometric empirical evidence on the macroeconomic rebound effects. The first paper that presents econometric evidence on macroeconomic rebound effects seems to be Orea et al. (2015), where the authors propose an energy demand stochastic frontier model with non-zero rebound effects. The authors apply the model in an empirical study using data on US states. They provide evidence of considerable high-energy rebound effects. Adetutu et al. (2016) assess the rebound effects for a panel of 55 countries. In their specification, they employ a two-stage approach, wherein in the first stage, an input distance function is utilized for measuring the country-level energy efficiency scores. In the second stage, they utilize a dynamic panel model where they estimate the effect of energy efficiency on aggregate energy consumption. Their method allows obtaining both short-run and long-run rebound effects. Last, using time series data (in monthly and quarterly frequency), Bruns et al. (2021) and Berner et al. (2022) utilize Structural and Factor vector autoregressive models (SVAR & FAVAR), respectively, and provide evidence for high partial rebound effects for several European economies and the USA. In the same spirit, Jafari et al. (2022) utilize an SVAR model for studying the Iranian economy, illustrating that the rebound effects are approximately 84%.

Our econometric approach is closely related to the method proposed by Orea et al. (2015), where the authors proposed an energy demand stochastic frontier approach with an inefficiency correction term, to allow non-proportional effects between the inefficiency component and the dependent variable. The authors, restrict the magnitude of the rebound effect on the $[0, 1)$ interval allowing only partial rebound effects to be present. To accomplish identification, they utilize the logistic function and express the rebound effects as a function of different covariates. As explained above, even though our specification does not allow for time-varying rebound effects, the major advantage of our approach is that it suggests a flexible way to identify all sorts of rebound effects, such as negative and zero rebound effects. In addition, in contrast to Adetutu et al. (2016), where the authors use an input distance function and measure the feasible contraction of all production inputs, we utilize an energy input distance function, and our estimated distance can be interpreted as the potential contraction of energy input conditional on all other inputs. Moreover, our study seems to be one of the few empirical studies (Adetutu et al. (2016) seems to be the only study to present country-level estimates) that present country-level rebound effects using annual data.

3.2 Economy-wide Rebound Effects

To begin with, in this section, we present a brief explanation and definition of the energy efficiency rebound effects. According to the energy economics literature, the rebound effect is defined as the percentage of potential savings that were not achieved due to several macroeconomic factors. A simple equation of the definition of the energy rebound effects is:

$$R = \left[1 - \frac{\text{Actual Savings}}{\text{Potential Savings}}\right] \times 100\%$$

where Actual Savings is the observed energy consumption decrease during a specific period of time, and Potential Savings are the savings that we would expect from an energy-efficient action or investment that is implemented. For instance, if an agent implements an energy efficiency investment which, from an engineering point of view, is expected to achieve 40% energy savings and the observed savings were 30%, we say that the rebound effect is $R = 1 - 0.3/0.4 = 0.25$ or 25%.

Therefore, we define the economy-wide energy rebound effect using the standard definition that appears in the literature (e.g. Saunders (2000), Orea et al. (2015), among others), as:

$$R = 1 - \eta^f \tag{3.1}$$

where η^f denotes the elasticity of energy consumption E with respect to energy efficiency improvements. In the next table, we present all the different potential rebound effect outcomes and the corresponding energy efficiency elasticities.

Table 3.1: Economy-wide Rebound Effect Scenarios

1	$R > 1$	Backfire	$\eta^f > 0$
2	$R = 1$	Full Rebound	$\eta^f = 0$
3	$0 < R < 1$	Partial Rebound	$-1 < \eta^f < 0$
4	$R = 0$	Zero Rebound	$\eta^f = -1$
5	$R < 0$	Super-Conservation	$\eta^f < -1$

Hence to draw inferences regarding the macroeconomic rebound effects, the estimation of the elasticity between energy efficiency and aggregate energy consumption, viz. η^f, is required. In the next section, we present our identification strategy.

3.3 Identification Strategy

3.3.1 Definition of Energy Efficiency

We consider a production process where each economy is using capital, labour and energy inputs $\{K, L, E\} \in \mathbb{R}^{+}$, to produce output $Y \in \mathbb{R}^{+}$. This production technology can be written as:

$$T(t) = \{(K, L, E, Y) : K, L, E \text{ can produces } Y \text{ at time } t\} \quad , \quad t = 1, 2, \dots, T \tag{3.2}$$

In this paper, to identify the economy-wide energy efficiencies, we follow recent literature developments based on Boyd (2008), Stern (2012), Filippini and Hunt (2011, 2015), and Tajudeen (2021), among others, and we define the economy-wide energy efficiency in a similar manner as:

$$EF = \frac{E^{opt}}{E} \in (0, 1] \tag{3.3}$$

where E^{opt} denotes the minimum/optimum energy consumption needed to produce the required services and products in an economy, and E is the measured/actual energy consumption in the economy. Therefore, from equation 3.3, we see that the level of energy efficiency indicates the deviations of the observed energy consumption from the minimum feasible one. Specifically, when an economy consumes higher energy units than the minimum energy required, viz. $E^{opt} < E$, we can say that the economy is not fully energy efficient; there are energy wastes and, as a result, the underlying energy efficiency will be $EF < 1$. On the other hand, when the economy operates using the minimum feasible units of energy, viz. $E^{opt} = E$, the market is fully energy efficiency $(EF = 1)$ and there are no energy wastes.

That said, from the definition of energy efficiency in 3.3, it can be seen that the energy efficiency estimation requires the knowledge of the latent minimum energy consumption. A common way to address this issue is to parameterize the latent minimum energy use using a parametric function of output and production inputs as:

$$\frac{h(Y, K, L)}{E} = exp(-u) \quad , \quad u \geq 0 \tag{3.4}$$

where $h(Y, K, L)$ usually takes the form of a Cobb-Douglas or any other flexible functional form, and the term $exp(-u)$ denotes the energy efficiency, which by definition should lie on the interval $(0, 1]$. Hence, to model energy efficiency, we utilize stochastic frontier analysis as an econometric tool to measure the energy inefficiencies or, in simpler words, to measure the distance of the observed energy consumption from the "hypothetical" minimum feasible energy use.

Applying the natural logarithm from both sides in equation 3.4, and re-arranging we have:

$$e = TL(y, k, l) + u \tag{3.5}$$

where e is the natural logarithm of E, $TL(y, k, l)$ denotes a translog function of log-output y, log-capital k and log-labour l and $u \geq 0$ denotes the energy inefficiency. Equation 3.5 is equivalent to an energy input distance function or energy demand stochastic frontier model as utilized by Boyd (2008), Stern (2012), Filippini and Hunt (2011, 2015), Tajudeen (2021), among others.

Having defined the above energy input distance function, the final form of our model is:

$$\begin{aligned} e_{it} =& \beta_k k_{it} + \beta_l l_{it} + \frac{1}{2}\beta_{kk} k_{it}^2 + \frac{1}{2}\beta_{ll} l_{it}^2 + \beta_{kl} k_{it} l_{it} \\ &+ \beta_y y_{it} + \frac{1}{2}\beta_{yy} y_{it}^2 + \beta_{yk} y_{it} k_{it} + \beta_{yl} y_{it} l_{it} + v_{it} + u_{it} \end{aligned} \tag{3.6}$$

where u_{it} denotes the energy inefficiency of country i at time t and v_{it} is the usual two-sided error term. The empirical model has a stochastic frontier econometric representation and can be written more compactly as:

$$e_{it} = \boldsymbol{z}'_{it}\boldsymbol{\beta} + v_{it} + u_{it} \tag{3.7}$$

where $\boldsymbol{z}_{it}$ is a vector including all the variables io the right-hand side of equation 3.6, including the η_i and η_t that are the country and time effects, respectively, $v_{it} \sim N(0, \sigma_v^2)$ is the usual two-sided error term and $u_{it} \geq 0$ is the non-negative inefficiency term. In this paper, we follow Aigner et al. (1977) and assume that the inefficiency term follows $u_{it} \sim N^+(0, \sigma_u^2)$.

According to equation 3.7, the usual stochastic frontier specification suggests a model with zero-rebound effects. In particular, the elasticity of aggregate energy demand with respect to inefficiency is:

$$\frac{\partial e_{it}}{\partial u_{it}} = 1 \qquad \text{and} \qquad R = 1 - \frac{\partial e_{it}}{\partial u_{it}} = 0$$

which implies that energy efficiency improvements will have a proportional effect on aggregate energy demand. Thus in this form, the identification of rebound effects is not possible, and the model implies zero rebound effects. Therefore, to resolve the issue of the energy-rebound identification, we proceed according to Orea et al. (2015), and we introduce a country specific "correction factor" $C_i = (1 - R_i)$, that interacts with the time-varying inefficiency u_{it}. The empirical model will have the form:

$$e_{it} = \boldsymbol{z}'_{it}\boldsymbol{\beta} + v_{it} + C_i u_{it} \tag{3.8}$$

Thus, incorporating the above correction allows the model to identify all sorts of energy rebound effects (values for $R_i < 1$), since the elasticity of energy inefficiency, is given by:

$$\frac{\partial e_{it}}{\partial u_{it}} = C_i = 1 - R_i$$

Here we need to highlight the fact, that under a frontier specification, the model will be well-defined only for $R_i < 1$. It is obvious that for rebound effects with $R_i \geq 1$, the model implies the inefficiency term to be $u_{it} \leq 0$, which is impossible in our framework. Thus, as Orea et al. (2015) note, if there is a-prior belief for backfire effects, frontier-based methods can not be appropriate tools for identification.

Consequently, the specification in 3.8, implies that the composed inefficiency $\tilde{u}_{it} = C_i u_{it}$ follows a half-normal distribution with country-specific variance σ^2_{ui}. This follows from the first and the second moments (mean and variance) of the half-normal distribution, where:

$$E(C_i u_{it}) = C_i E(u_{it}) = \sigma_{ui} \frac{\sqrt{2}}{\sqrt{\pi}}$$
$$Var(C_i u_{it}) = C_i^2 Var(u_{it}) = C_i^2 \sigma_u^2 \left(1 - \frac{2}{\pi}\right) = \sigma^2_{ui} \left(1 - \frac{2}{\pi}\right)$$

Hence, the energy input distance function with a country-specific energy rebound effect implies a stochastic frontier specification with country-specific heteroscedasticity.

3.3.2 The Econometric Model

The reduced form econometric specification will have the form:

$$\begin{gathered} e_{it} = \boldsymbol{z}'_{it}\boldsymbol{\beta} + v_{it} + \tilde{u}_{it} \\ v_{it} \sim \mathcal{N}(0, \sigma_v^2) \quad , \quad \tilde{u}_{it} \sim \mathcal{N}^+(0, \sigma^2_{ui}) \\ i = 1, 2, \ldots, N \quad , \quad t = 1, 2, \ldots, T \end{gathered} \tag{3.9}$$

From the above model, we can see that the structural parameters to be estimated is the vector $\boldsymbol{\theta} = [\boldsymbol{\beta}, \sigma_v, \{\sigma_{ui}\}_{i=1,2,\ldots,N}]'$. Additionally, we assume that the pdf of the vector $\boldsymbol{z}_{it}$ is independent of the v_{it} and u_{it} and is not evolving any of the structural parameters $\boldsymbol{\theta}$. Hence, the joint density function will be given by:

$$f(e_{it}, \tilde{u}_{it} | \boldsymbol{z}_{it}, \boldsymbol{\theta}) \propto f_N(e_{it} | \boldsymbol{z}'_{it}\boldsymbol{\beta} + \tilde{u}_{it}, \sigma_v^2) f_N(\tilde{u}_{it} | 0, \sigma^2_{ui}) \boldsymbol{I}(u_{it} \geq 0) \tag{3.10}$$

for each $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$, where $f_N(.|A, B)$ denotes the pdf of the normal distribution with mean A and variance B. However, the composed inefficiency term $\tilde{u}_{it}$ is a latent term, and as a result, to able to proceed with the model estimation, the inefficiency term $\tilde{u}_{it}$ has to be integrated out of the joint pdf in 3.10. Hence, the marginal density can be obtained by the following integration:

$$f(e_{it} | \boldsymbol{z}_{it}, \boldsymbol{\theta}) = \int_{\tilde{u}_{it} \in \mathcal{R}_+} f_N(e_{it} | \boldsymbol{z}'_{it}\boldsymbol{\beta} + \tilde{u}_{it}, \sigma_v^2) f_N(\tilde{u}_{it} | 0, \sigma^2_{ui}) \boldsymbol{I}(\tilde{u}_{it} \geq 0) d\tilde{u}_{it} \tag{3.11}$$

The solution of the above integral has been presented in Aigner et al. (1977) and is available in closed form:

$$f(e_{it} | \boldsymbol{z}_{it}, \boldsymbol{\theta}) = \frac{2}{\sigma_i} \phi\left(\frac{\varepsilon_{it}}{\sigma_i}\right) \Phi\left(\frac{\lambda_i}{\sigma_i} \varepsilon_{it}\right) \tag{3.12}$$

where $\varepsilon_{it} = e_{it} - \boldsymbol{z}'_{it}\boldsymbol{\beta}$, $\lambda_i = \sigma_{ui}/\sigma_v$ and $\sigma_i^2 = \sigma_v^2 + \sigma_{ui}^2$. Last, the likelihood function

$$L(\boldsymbol{e}|\boldsymbol{Z}, \boldsymbol{\theta}) = \prod_{i=1}^{N}\prod_{t=1}^{T} f(e_{it}|\boldsymbol{z}_{it}, \boldsymbol{\theta}) \tag{3.13}$$

can be maximized with respect to the unknown structural parameters $\boldsymbol{\theta} = [\boldsymbol{\beta}, \sigma_v, \{\sigma_{ui}\}_{i=1,2,\ldots,N}]'$. However, as we can see from equations 3.10-3.11, the likelihood function $L(\boldsymbol{e}|\boldsymbol{Z}, \boldsymbol{\theta})$ is highly non-linear on the parameters $\boldsymbol{\theta} = [\boldsymbol{\beta}, \sigma_v, \{\sigma_{ui}\}_{i=1,2,\ldots,N}]'$, and the maximization of the log-likelihood usually involves complex numerical estimation techniques.

For this reason, we proceed with Bayesian analysis (see Van den Broeck et al. (1994), Fernandez et al. (1997)) using Markov Chain Monte Carlo methods where the latent inefficiencies are considered as parameters to be estimated. This method is called data-augmentation and has been initially introduced by Tanner and Wong (1987)[1].

Following the Bayes rule, we have:

$$p(\boldsymbol{\theta}, \tilde{\boldsymbol{u}}|\boldsymbol{e}, \boldsymbol{Z}) \propto f(\boldsymbol{e}|\boldsymbol{Z}, \tilde{\boldsymbol{u}}, \boldsymbol{\theta}) f(\tilde{\boldsymbol{u}}|\boldsymbol{\theta}) \pi(\boldsymbol{\theta}) \tag{3.14}$$

where $p(\boldsymbol{\theta}, \tilde{\boldsymbol{u}}|\boldsymbol{e}, \boldsymbol{Z})$ is the posterior distribution of the unknown structural parameters $\boldsymbol{\theta}$ and the composed inefficiency vector $\tilde{\boldsymbol{u}}$, $f(\boldsymbol{e}|\boldsymbol{Z}, \tilde{\boldsymbol{u}}, \boldsymbol{\theta})$ is the likelihood function conditioning on the unknown structural parameters $\boldsymbol{\theta}$ and the latent inefficiency vector $\tilde{\boldsymbol{u}}$, $f(\tilde{\boldsymbol{u}}|\boldsymbol{\theta})$ denotes the half-normal pdf of the latent composed inefficiency vector and $\pi(\boldsymbol{\theta})$ is the prior distribution of the structural parameters. Before proceeding with the Bayesian estimation, one needs to define the prior distributions of the parameters:

$$\pi(\boldsymbol{\theta}) \propto \pi(\boldsymbol{\beta})\pi(\sigma_v^2)\prod_{i=1}^{n}\pi(\sigma_{ui}^2) \tag{3.15}$$

For the parameter vector $\boldsymbol{\beta}$ we assume a non-informative flat prior as:

$$\pi(\boldsymbol{\beta}) \propto 1 \tag{3.16}$$

For the variances of the error term and the inefficiency term, we use the Inverse-Gamma distribution:

$$\sigma_v^2|a_0, a_1 \sim \mathcal{IG}(a_0, a_1) \quad , \quad \sigma_{ui}^2|q_0, q_1 \sim \mathcal{IG}(q_0, q_1) \quad \text{for} \quad i = 1, 2, \ldots, n \tag{3.17}$$

where a_0, q_0 and a_1, q_1 are the corresponding shape and scale parameters, respectively. The choice of the Inverse-Gamma distribution is a standard approach in Bayesian econometrics since it is the natural conjugate prior distribution for the variance of a normal distribution. For the hyperparameters of σ_v we choose $a_0 = a_1 = 10^{-2}$, and for all σ_{ui}'s we set $q_0 = q_1 = 10^{-2}$.

[1] Tsionas (2001) presents an excellent simulation study for the Bayesian stochastic frontier model with truncated normal distribution.

The choice of these hyperparameters is standard in bayesian analysis and produces the non-informative Jeffrey's priors.

Given the above prior distributions for the parameter vector $\boldsymbol{\Theta} = [\boldsymbol{\beta}, \sigma_v, \{\sigma_{ui}\}_{i=1,2,\dots,N}, \tilde{\boldsymbol{u}}]'$, the joint augmented posterior distribution will have the form:

$$
\begin{aligned}
p(\boldsymbol{\Theta}|\boldsymbol{e}, \boldsymbol{Z}) \propto (\sigma_v^2)^{-\frac{NT}{2}} exp\left\{-\frac{(\boldsymbol{e}-\boldsymbol{Z\beta}-\tilde{\boldsymbol{u}})'(\boldsymbol{e}-\boldsymbol{Z\beta}-\tilde{\boldsymbol{u}})}{2\sigma_v^2}\right\} \\
\times \prod_{i=1}^{N}\left[(\sigma_{ui}^2)^{-\frac{T}{2}} exp\left\{-\frac{\tilde{\boldsymbol{u}}_{\boldsymbol{i}}{}'\tilde{\boldsymbol{u}}_i}{2\sigma_{ui}^2}\right\} \boldsymbol{I}(\tilde{u}_{it} \geq 0)\right] \\
\times (\sigma_v^2)^{-a_0-1} exp\left\{-\frac{a_1}{\sigma_v^2}\right\} \\
\times \prod_{i=1}^{N}\left[(\sigma_{ui}^2)^{-q_0-1} exp\left\{-\frac{q_1}{\sigma_{ui}^2}\right\}\right]
\end{aligned}
\tag{3.18}
$$

where the first line denotes the likelihood function, the second line illustrates the half-normal distributions of the inefficiency vector $\tilde{\boldsymbol{u}}_i$ and the third and fourth lines, depict the inverse-gamma prior distribution of σ_v and σ_{ui}'s. From the above, we are able to obtain the corresponding conditional distributions for the unknown parameter vector $\boldsymbol{\Theta}$. More specifically, we have:

- Conditional distribution of $\boldsymbol{\beta}$:

$$
\boldsymbol{\beta}|\boldsymbol{\Theta}_{-\boldsymbol{\beta}} \sim \mathcal{N}((\boldsymbol{Z}'\boldsymbol{Z})^{-1}(\boldsymbol{Z}'(\boldsymbol{e}-\tilde{\boldsymbol{u}})), \sigma_v^2(\boldsymbol{Z}'\boldsymbol{Z})^{-1}) \tag{3.19}
$$

- Conditional distribution of σ_v^2:

$$
\sigma_v^2|\boldsymbol{\Theta}_{-\boldsymbol{\sigma_v^2}} \sim \mathcal{IG}\left(\frac{NT}{2} + a_0, \frac{(\boldsymbol{e}-\boldsymbol{Z\beta}-\tilde{\boldsymbol{u}})'(\boldsymbol{e}-\boldsymbol{Z\beta}-\tilde{\boldsymbol{u}})}{2} + a_1\right) \tag{3.20}
$$

- Conditional distribution of σ_{ui}^2 for $i = 1, 2, \dots, N$:

$$
\sigma_{ui}^2|\boldsymbol{\Theta}_{-\boldsymbol{\sigma_{ui}^2}} \sim \mathcal{IG}\left(\frac{T}{2} + q_0, \frac{\tilde{\boldsymbol{u}}_i'\tilde{\boldsymbol{u}}_i}{2} + q_1\right) \tag{3.21}
$$

- Conditional distribution of latent inefficiency vector $\tilde{\boldsymbol{u}}_i$ for $i = 1, 2, \dots, N$:

$$
\tilde{\boldsymbol{u}}_i|\boldsymbol{\Theta}_{-\tilde{\boldsymbol{u}}_i} \sim \mathcal{N}^+\left(\frac{\sigma_{ui}^2(\boldsymbol{e}_i - \boldsymbol{Z}_i\boldsymbol{\beta})}{\sigma_{ui}^2 + \sigma_v^2}, \frac{\sigma_v^2\sigma_{ui}^2}{\sigma_{ui}^2 + \sigma_v^2}\right) \tag{3.22}
$$

Once the conditional distributions are known, we can proceed with posterior inference using different MCMC techniques, for instance, the Gibbs sampler algorithm. The Gibbs Sampler helps us to obtain a sequence of posterior samples which enables us to approximate the marginal posterior distribution of the corresponding unknown parameters. To obtain point estimates for all the parameters, one can use the posterior average or the posterior median.

3.3.3 The Rebound Effects

Once we obtain the estimates from the reduced form stochastic frontier model, we can proceed with the identification of the unknown rebound effects R_i's and the σ_u. To identify these parameters of interest, we use the theoretical model conditions which follow from the mean and the variance of the half-normal distribution, along with the information we obtained from the reduced form stochastic frontier model regarding the composed inefficiency term $\tilde{u}_{it}$ and the country-specific σ_{ui}'s.

The first set of moment conditions consists of the following N conditions:

$$(1 - R_i)\sigma_u - \hat{\sigma}_{ui} = 0 \qquad \text{for} \qquad i = 1, 2, \dots, N \tag{3.23}$$

Moreover, we utilise the first and the second moments of the half-normal distribution. Specifically, we have:

$$(1 - R_i)\sigma_u \frac{\sqrt{2}}{\sqrt{\pi}} - E(\hat{\tilde{\boldsymbol{u}}}_i) = 0 \qquad \text{for} \qquad i = 1, 2, \dots, N \tag{3.24}$$

and

$$(1 - R_i)^2 \sigma_u^2 \left(1 - \frac{2}{\pi}\right) - Var(\hat{\tilde{\boldsymbol{u}}}_i) = 0 \qquad \text{for} \qquad i = 1, 2, \dots, N \tag{3.25}$$

Therefore, we are able to elaborate the $3 \times N$ moment conditions in 3.23, 3.24 and 3.25 to identify the $N+1$ unknown parameter vector $\boldsymbol{\Lambda} = [R_1, R_2, \dots, R_N, \sigma_u]'$. Of course, one can include additional moment conditions by including the squared of the aforementioned moments, etc. To identify the parameter vector $\boldsymbol{\Lambda}$ that achieves the matching of the above moment conditions, one can use standard econometric tools such as the Generalized Method of Moments (GMM) (Hansen (1982)). Thus, in this step, the posterior estimates regarding the $\hat{\tilde{u}}_{it}$ and $\hat{\sigma}_{ui}$, will be considered as known and they will be treated as observed data, viz. $\mathcal{F} = [\hat{\tilde{\boldsymbol{u}}}_1, \hat{\tilde{\boldsymbol{u}}}_2, \dots, \hat{\tilde{\boldsymbol{u}}}_N, \{\hat{\sigma}_{ui}\}_{i=1,2,\dots,N}]'$.

More compactly, the moment conditions $g(\mathbf{\Lambda}|\mathcal{F})$ can be written in a vector form as:

$$g(\mathbf{\Lambda}|\mathcal{F}) = \mathbb{E}\begin{bmatrix} (1-R_1)\sigma_u - \hat{\sigma}_{u1} \\ (1-R_2)\sigma_u - \hat{\sigma}_{u2} \\ \vdots \\ (1-R_N)\sigma_u - \hat{\sigma}_{uN} \\ (1-R_1)\sigma_u \frac{\sqrt{2}}{\sqrt{\pi}} - E(\hat{\hat{\boldsymbol{u}}}_1) \\ (1-R_2)\sigma_u \frac{\sqrt{2}}{\sqrt{\pi}} - E(\hat{\hat{\boldsymbol{u}}}_2) \\ \vdots \\ (1-R_N)\sigma_u \frac{\sqrt{2}}{\sqrt{\pi}} - E(\hat{\hat{\boldsymbol{u}}}_N) \\ (1-R_1)^2\sigma_u^2\left(1-\frac{2}{\pi}\right) - Var(\hat{\hat{\boldsymbol{u}}}_1) \\ (1-R_2)^2\sigma_u^2\left(1-\frac{2}{\pi}\right) - Var(\hat{\hat{\boldsymbol{u}}}_2) \\ \vdots \\ (1-R_N)^2\sigma_u^2\left(1-\frac{2}{\pi}\right) - Var(\hat{\hat{\boldsymbol{u}}}_N) \end{bmatrix} = \mathbf{0} \tag{3.26}$$

where $g(\mathbf{\Lambda}|\mathcal{F})$ corresponds to the moment vector. Here we should highlight that we can replace the $E(\hat{\hat{\boldsymbol{u}}}_i)$ and $Var(\hat{\hat{\boldsymbol{u}}}_i)$ for $i = 1, 2, \ldots, N$, with their corresponding sample moments.

To estimate the unknown parameters, the GMM seeks to minimize the criterion below:

$$Q(\mathbf{\Lambda}) = arg\ min\ g(\mathbf{\Lambda}|\mathcal{F})'\boldsymbol{W} g(\mathbf{\Lambda}|\mathcal{F}) \tag{3.27}$$

where $\boldsymbol{W}$ is a $(L \times L)$ positive definite weighting matrix, with $L = 3 \times N$ is the number of moment conditions. It is well known that the GMM estimator, under well-defined moment conditions, is consistent and asymptotically normal.

For the first step, a common choice is to use the identity matrix ($\boldsymbol{W} = \boldsymbol{I}$) by giving the same weights to all moment conditions. This is known as the 1-step GMM (GMM-1), and we can obtain consistent estimates that are asymptotically normal. Despite its simplicity, it is known that this estimator is not efficient. For this reason, given the estimates from the GMM-1, we can proceed with the second stage (GMM-2), where the GMM minimizes the criterion in 3.27 using the $\hat{\boldsymbol{W}} \propto [g(\hat{\mathbf{\Lambda}}_1|\mathcal{F})'g(\hat{\mathbf{\Lambda}}_1|\mathcal{F})]^{-1}$, where $\hat{\mathbf{\Lambda}}_1$ are the GMM-1 estimates. The intuition behind the GMM-2 is that we update the weights according to the importance of each moment condition. The GMM-2 is well known to be asymptotically more efficient than the GMM-1 estimator[2].

[2]Similar estimators include the iterated (GMM-Iterated) and the continuous updating GMM (GMM-CUE).

3.4 Empirical Study

3.4.1 Energy Efficiency estimates and the Rebound Effects

Here we implement the proposed econometric identification procedure in an empirical example where we estimate the energy efficiency scores and the corresponding rebound effects for several developed economies. We use a balanced panel of 20 OECD member states[3] using data on Output, Capital, Labour and Energy from 1980 to 2018. The choice for the number of countries and the period is driven by data availability. In Table 3.2, we present descriptive statistics of the variables used in our study.

Table 3.2: Descriptive Statistics of the variables

Variable	Units	Mean	Std	Source
Output	GDP (in mil. 2017US$, PPP)	1622163	3031095	PWT 10.0
Capital	Capital Stock (in min. 2017US$, PPP)	6918161	11462996	PWT 10.0
Labour	Persons Engaged (in mil.)	19.60	30.37	PWT 10.0
Energy	TEC (in quad Btu)	9.53	19.50	U.S. EIA
Industrial Production	Index (2015=100)	85.63	23.59	OECD
Energy Prices	Index (2015=100)	67.62	26.13	OCED Database
Trade Openness	Exports + Imports (% of GDP)	0.79	0.45	PWT 10.0
TFP	Solow Residual	1107.02	257.09	Own Calculation
FDI	Weighted Index $\in [0, 1]$	0.62	0.20	IMF Database

Moreover, we control for a number of macroeconomic variables that could possibly explain the aggregate energy demand and lessen problems arising from the omitted variable bias. In particular, we include the Industrial Production Index, which can be used as a proxy for the energy intensity of each economy, the Energy Prices, the Trade Openness, which captures the energy demand embedded in trade activity, the Total Factor Productivity reflecting the technological progress in each economy and last, the Financial Development Index which developed by the International Monetary Fund (IMF) and reflects the ability of domestic firms to have access to credit and therefore financing investments. In addition, we control for common time effects using time dummies. Moreover, to control for the unobserved heterogeneity across the countries, we estimate the True-Fixed Effects (TFE) stochastic frontier model as presented by Greene (2005a,b). As Greene reported, failing to control for country-specific effects will generate biased energy efficiency estimates[4].

[3]The countries included in our study are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, UK and USA.

[4]The importance of accounting for the unobserved heterogeneity effects has been extensively elaborated in

Table 3.3: Posterior Parameter Estimates

	TFE-SFM	
	Post. Mean	Post. Std
Frontier Parameters		
$ln(K)$	0.1377	0.0282
$ln(L)$	0.7077	0.0644
$0.5\ ln(K)^2$	-0.1191	0.0720
$0.5\ ln(L)^2$	-0.5734	0.0724
$ln(K)\ ln(L)$	0.3654	0.0488
$ln(Y)$	0.0365	0.0534
$0.5\ ln(Y)^2$	-0.1098	0.1094
$ln(K)\ ln(Y)$	-0.2609	0.0756
$ln(L)\ ln(Y)$	0.3446	0.0699
Control Variables		
$ln(Price)$	0.0001	0.0174
$ln(Ind)$	0.0769	0.0189
$ln(TradeOP)$	0.0119	0.0189
$ln(Fin)$	0.1088	0.0133
Country Effects	Yes	
Time Dummies	Yes	
σ_v	0.0231	0.0016
Composed Inefficiency Term		
Median	0.0460	
Average	0.0743	
Max	0.5034	
Min	0.0052	

Note: Posterior analysis has been performed using 30,000 iterations from which the first 10,000 draws have been discarded to eliminate any initial values effect. In our model, we exclude the TFP variable since it creates collinearities issues. In addition, we have estimated the model, including the TFP and excluding the time effects; and the results are virtually the same.

In Table 3.3, we report the posterior parameter estimates and descriptive statistics of the corresponding composed inefficiency term. From the TFE-SFM presented in Table 3.3, we see that all the elasticities seem to have reasonable values. In particular, the average capital, labour and income elasticity have positive values of 0.19, 0.51 and 0.16, respectively. Regarding the elasticity of industrial production, the posterior average is 0.0769 (posterior std is 0.0189), indicating that energy-intensive economies require higher energy use. On the other hand, we find that trade openness does not play a crucial role in determining aggregate energy demand.

the stochastic frontier literature. Some recent studies that elaborate panel stochastic frontier models under fixed effects consist of Wang and Ho (2010), Chen et al. (2014), Belotti and Ilardi (2018), Kutlu et al. (2019), among others.

Table 3.4: Energy Efficiency Rebound Effects estimates

Country	Parameter	GMM-1	GMM-2
Australia	$\hat{R}_1$	0.5649	0.5407
Austria	$\hat{R}_2$	0.4231	0.4002
Belgium	$\hat{R}_3$	0.6632	0.6716
Canada	$\hat{R}_4$	0.6930	0.7236
Denmark	$\hat{R}_5$	-0.0301	-0.0445
Finland	$\hat{R}_6$	0.7202	0.7878
France	$\hat{R}_7$	0.7670	0.9241
Germany	$\hat{R}_8$	-0.1975	-0.2005
Ireland	$\hat{R}_9$	0.4647	0.4331
Italy	$\hat{R}_{10}$	0.7504	0.8649
Japan	$\hat{R}_{11}$	0.7426	0.8536
Netherlands	$\hat{R}_{12}$	0.7375	0.8332
New Zealand	$\hat{R}_{13}$	0.5088	0.4767
Norway	$\hat{R}_{14}$	0.5737	0.5481
Portugal	$\hat{R}_{15}$	0.5146	0.4743
Spain	$\hat{R}_{16}$	0.6535	0.6541
Sweden	$\hat{R}_{17}$	0.3066	0.2759
Switzerland	$\hat{R}_{18}$	0.7311	0.8248
UK	$\hat{R}_{19}$	-0.7867	-0.8099
USA	$\hat{R}_{20}$	0.6077	0.5909
	$\hat{\sigma}_u$	0.1802	0.1782

Note: The table reports the parameters estimates using the 1-stage GMM and the 2-stage GMM. We have used $(6 \times N)$ to identify $N + 1$ parameters. To minimize the GMM criterion, we have used the quasi-Newton algorithm. All computations have been performed in the R programming language.

Interestingly, the energy prices seem not to affect aggregate energy demand, indicating that the aggregate energy demand is inelastic in energy price shocks. Similar results have been reported in various energy economics-related studies. Last, the estimated standard deviation of the error term seems to be very low (the posterior average is 0.0231), indicating that the proposed specification fits the data very well.

Next, we proceed with the estimation of the country-specific rebound effects and the corresponding energy efficiency scores. To do that, we use the proposed moment conditions to extract the relative information using the composed inefficiency term ($\hat{\hat{u}}_{it}$) and the country-specif $\hat{\sigma}_{ui}$'s obtained from the reduced form energy input stochastic distance function. In addition, we increase our moment conditions by including the squared of proposed orthogonality conditions ($3 \times N$ additional moments). In Table 3.4, we present the parameter estimates using the GMM-1 and GMM-2 methods.

From the above estimates, we find that during the last 40 years, most developed countries exhibited partial rebound effects ranging approximately from 28% to 92%. This shows that for 17 economies in our sample, part of the energy efficiency improvements has been offset by economic activity. In practical terms, this can be explained by the fact that part of the energy cost-effective improvements have been "re-invested" to support the high economic growth that has been observed during the last decades. These results are in line with recent econometric empirical findings, which highlight the fact that part of initial energy efficiency improvements will vanish (e.g. Orea et al. (2015), Adetutu et al. (2016), Bruns et al. (2021), Berner et al. (2022), among others).

Specifically, from Table 3.4, we see that the highest rebound effects are observed in France, Italy, Japan, Netherlands, Switzerland, Finland and Canada, ranging from approximately 72% to 92%. In the rest of the countries, we see modest rebound effects ranging from 40% to 67%. Moreover, Sweden seems to have the lowest partial rebound effect of 28%. On the other hand, Denmark is the only country with a zero-rebound effect, suggesting that the economy has managed to capitalize effectively on all the energy efficiency cost-effective actions that have been implemented.

Interestingly, for Germany and the UK, the estimated rebound effect from the GMM-2 are -0.20 and -0.80, respectively, indicating that during the last decades, both economies exhibited an energy "super-conservation" behaviour. In practical terms, this suggests that during the last 40 years, both economies not only performed effective energy efficiency policies but also substantially reduced their energy-intensive economic activities. However, here we need to highlight that during the 1980s and 1990s, both economies underwent a series of significant economic reforms. For instance, in the United Kingdom, the government implemented significant economic reforms aiming at modernizing the economy and significantly increasing competition and efficiency in the markets. Specifically, many state-owned firms were privatised, leading to significantly higher financial performance and cost-effective actions (Cragg and Dyck (1999)). Moreover, we believe that other reasons that contributed to the super-conservation effect are behavioural changes related to environmental awareness, such as the increasing use of public transportation and others.

Regarding Germany, we note that the rebound effect estimates can be affected due to historical events that happened during the late 80s and early 90s, when West and East Germany reunified, and consequently, we believe that part of the energy super-conservation effect should be attributed to this period. Specifically, from the raw data, we can see a significant break point of the aggregate energy use, where in 1990 the energy use was 16.80 BTU, and dropped to 14.96 BTU in 1991, indicating a huge decline of approximately 11% in a single year.

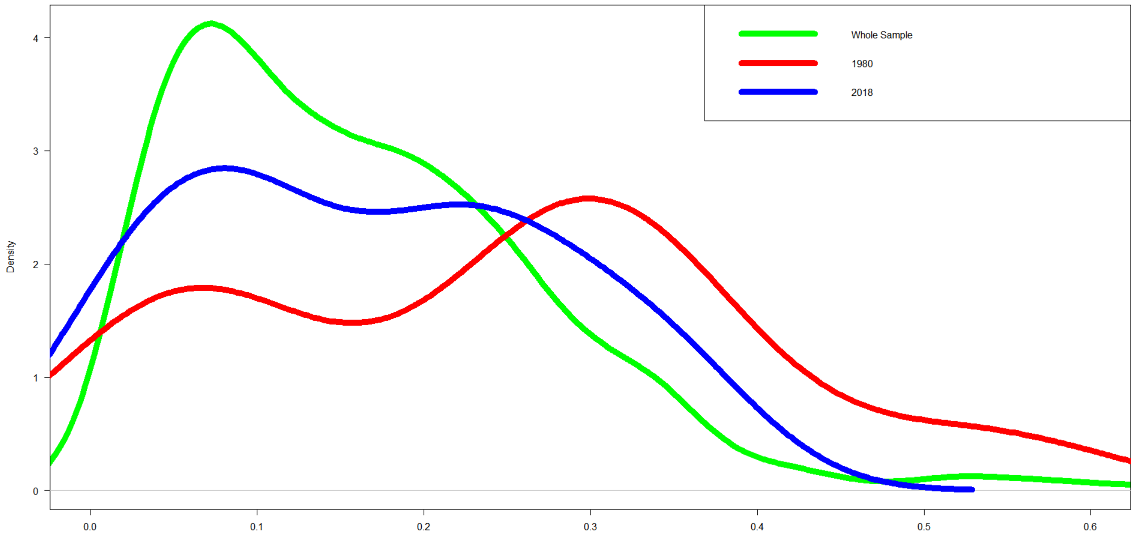


Figure 3.1: Energy Inefficiency density

For this reason, to assess the robustness of our analysis, in the next subsection, we present some robustness exercises where we re-estimate the model and the rebound effects restricting our sample period for the last decade of our sample (2008-2018). The reason why we picked this time period is to match the time span with the study presented by Berner et al. (2022). For the perid 2008-2018, our results indicate partial rebound effects for all economies, including Germany and UK, indicating that the super-conservation effect is attributed to the arguments explained above.

Next, once we have obtained the estimates of the rebound effects $\hat{R}_i$ for $i = 1, 2, \ldots, N$, we can derive the energy inefficiency scores by:

$$u_{it} = \frac{\tilde{u}_{it}}{1 - \hat{R}_i} \qquad \text{for} \qquad i = 1, 2, \ldots, N$$

In Figure 3.1, we illustrate the energy inefficiency densities for the whole sample period and the years 1980 and 2018, respectively. Overall, we see that the blue density line illustrating the energy inefficiency during 2018 is shifted towards the zero inefficiency area, indicating a general upward trend for most of the economies involved in this study. In particular, the average and the median energy inefficiency for the year 2018 are both 0.169, respectively, suggesting that despite the energy efficiency improvements that have been observed, there is still room for further energy cost-effective savings. This is important in practice, especially during soaring energy costs and when achieving energy savings seems paramount.

Furthermore, in Table 3.5, we report the estimated energy efficiencies for the years 1980 and 2018, and the corresponding percentage change throughout our study. Table 3.5 illustrates

Table 3.5: Energy Efficiency Estimates by Country

Country	1980	2018	% Change	Country	1980	2018	% Change
Australia	97.5%	75.6%	-22.5%	Japan	92.5%	90.2%	-2.5%
Austria	59.8%	90.2%	+50.8%	Netherlands	69.8%	92.0%	+31.7%
Belgium	77.9%	81.6%	+4.7%	New Zealand	93.3%	89.9%	-3.6%
Canada	76.6%	70.5%	-8.0%	Norway	92.7%	79.8%	-13.9%
Denmark	70.5%	98.2%	+39.2%	Portugal	93.9%	82.8%	-11.8%
Finland	90.9%	79.0%	-13.1%	Spain	97.9%	77.9%	-20.5%
France	76.4%	71.2%	-6.8%	Sweden	74.5%	97.1%	+30.3%
Germany	72.1%	96.6%	+33.9%	Switzerland	70.1%	92.5%	+31.9%
Ireland	72.0%	70.1%	-2.7%	UK	77.5%	98.9%	+27.7%
Italy	55.0%	85.2%	+54.9%	USA	74.4%	76.7%	+3.2%

Note: The table illustrates the posterior median of the estimated energy efficiencies. In addition, in Appendix A and Figures 3.3 - 3.5, we illustrate the graphs of the intertemporal energy efficiency estimates.

that Austria, Denmark, Germany, Italy, Netherlands, Sweden, Switzerland, and the U.K. have managed to increase their energy efficiency by more than 20%. In addition, Belgium and USA seem to have improved their energy efficiency slightly, at a rate of 4.7% and 3.2%, respectively. On the other hand, Australia, Canada, Finland, Norway, Portugal and Spain, although they exhibited very high energy efficiency scores during the 80s, have not managed to maintain their energy efficiency performance during the following decades, and, their energy efficiency level has declined more than 10%. Last, Canada, France, Ireland, Japan and New Zealand exhibit a slight decrease in their energy use performance ranging from -2.5% to -6.8%.

Next, we assess the relationship between economy-wide rebound effects and each country's overall energy efficiency intertemporal trend and address whether energy-efficiency policies can be an effective long-run tool for achieving the green transition. In Figure 3.2, we present the relationship between the efficiency change (%) and the rebound effects.

Overall, we can see that the energy rebound effects have a heterogeneous impact on energy efficiency improvements. From Figure 3.2, we can see that we can cluster the countries into three groups. First, we see countries such as Austria, Belgium, Italy, Netherlands, Sweden, Switzerland and the USA, although their economies exhibit large rebound effects, achieved increased energy efficiency performance and benefit from cost-effective energy actions. Moreover, as mentioned above, Denmark, Germany and UK comprise the second group, with high positive energy efficiency change and super-conservation effect.

On the other hand, for the rest of the countries, viz. Australia, Canada, Finland, France, Ireland, Japan, New Zealand, Norway, Spain and Portugal, we see high macroeconomic rebound effects and declining intertemporal energy efficiency trends. One possible explanation is the fact that increasing energy efficiency performance or maintaining energy efficiency at high levels is

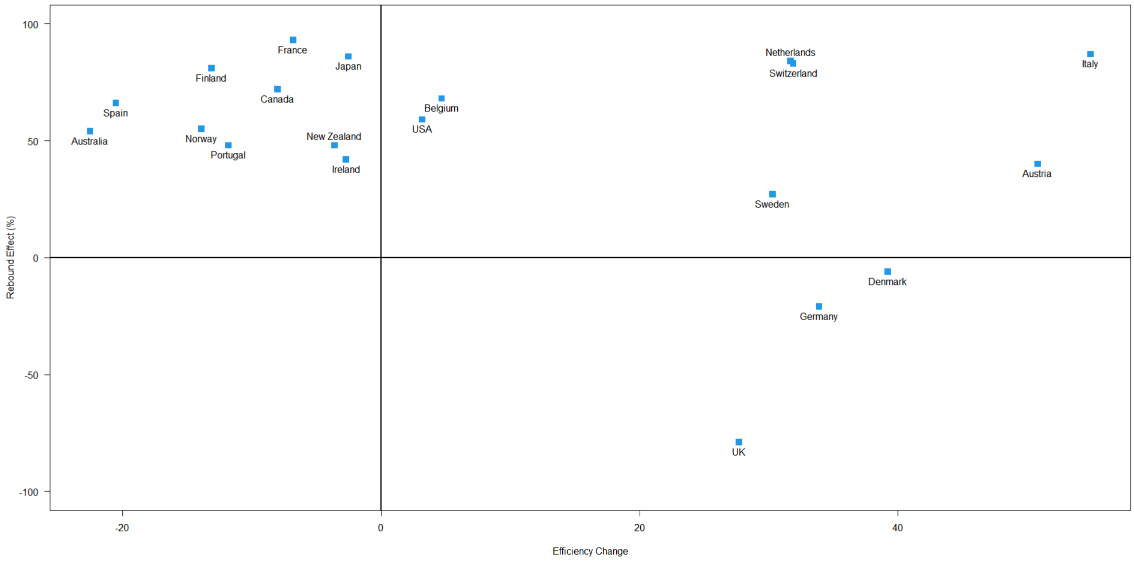


Figure 3.2: Efficiency Change and Rebound Effects (1980-2018)

costly (the third group of countries is characterized by high energy performance during the initial period), and as a result, the promotion of high economic growth required to sacrifice part of their energy performance. Moreover, for the majority of the period 1980 to 2018, energy prices remained at low and affordable levels, with very short periods of exuberance, and therefore, the boost of economic development using less energy-efficient technologies could be beneficial from a cost-benefit analysis point of view.

3.4.2 Robustness Tests for the Rebound Effects

In this section, we present the estimated rebound effects for the period 2008-2018. Specifically, we re-estimate the reduced form stochastic frontier model and re-estimate the rebound effects using the proposed moment conditions.

Overall, we see that for this specific sub-period, all countries exhibit partial rebound effects ranging from 27% to 70%. This result is consistent with most of the literature and re-confirms the robustness of the proposed method. For Denmark, Germany and the UK, the estimated rebound effects are 54%, 61% and 38%, respectively. This indicates that the zero-rebound effect and the super-conservation effect obtained in the initial analysis using the whole sample (Table 3.6, column 1), are driven mainly by the general industrial and market transitions introduced during the 80s and 90s. In addition, in line with the energy efficiency scores presented in Table 3.5, we observe that as energy efficiency improves, the economies are more willing to sacrifice energy savings to boost their economic activity using energy-intensive products and services.

Table 3.6: Rebound Estimates for the whole and restricted sample

	Full Sample	2008-2018
Australia	54%	53%
Austria	40%	58%
Belgium	67%	45%
Canada	72%	36%
Denmark	-4%	54%
Finland	79%	55%
France	92%	63%
Germany	-20%	61%
Ireland	43%	42%
Italy	86%	63%
Japan	85%	57%
Netherlands	83%	70%
New Zealand	48%	58%
Norway	55%	36%
Portugal	47%	27%
Spain	65%	44%
Sweden	28%	55%
Switzerland	82%	35%
UK	-81%	38%
USA	59%	62%

Moreover, our estimated magnitudes are lower but in line with the estimates obtained by Berner et al. (2022), where the authors report economy-wide rebound effects of 78% to 101% for France, Germany, Italy and the UK. We believe these differences should be mainly attributed to the difference between annual panel data analysis and time series analysis using monthly or quarterly data.

3.5 Conclusions and Policy Implications

This chapter provides a new method for estimating country-specific energy efficiency rebound effects. The main advantage of our model is that it treats the country-specific rebound effects as parameters to be estimated and does not impose any particular restriction regarding their magnitude. In addition, the proposed specification is relatively easy to be implemented in any standard statistical software.

Our empirical study reveals modest to considerable energy efficiency rebound effects for most OECD economies. In particular, the partial rebound effects range approximately from 28% to 92%, which supports the argument that energy efficiency actions could have a limited impact on achieving environmental objectives. Moreover, we provide evidence that the average energy efficiency for 2018 is approximately 84%, indicating that the OECD economies can further reduce their energy use without disrupting their economic activity. In addition, we find mixed evidence for the relationship between energy efficiency and the rebound effects.

In terms of current limitations and future extensions, we recognise that the assumption of country-specific rebound effects could be restricted in practice. For this reason, it will be of particular interest to allow the rebound effects to be time-varying and study the intertemporal evolution throughout time. Moreover, we plan to extend our dataset and include developing economies. This will enable us to conduct a comparative analysis regarding the energy efficiency trends and the magnitude of the rebound effects.

Appendix A

In this section, we illustrate the figures of the estimated energy efficiencies from 1980 to 2018. Each point in time reflects the median of the posterior distribution.

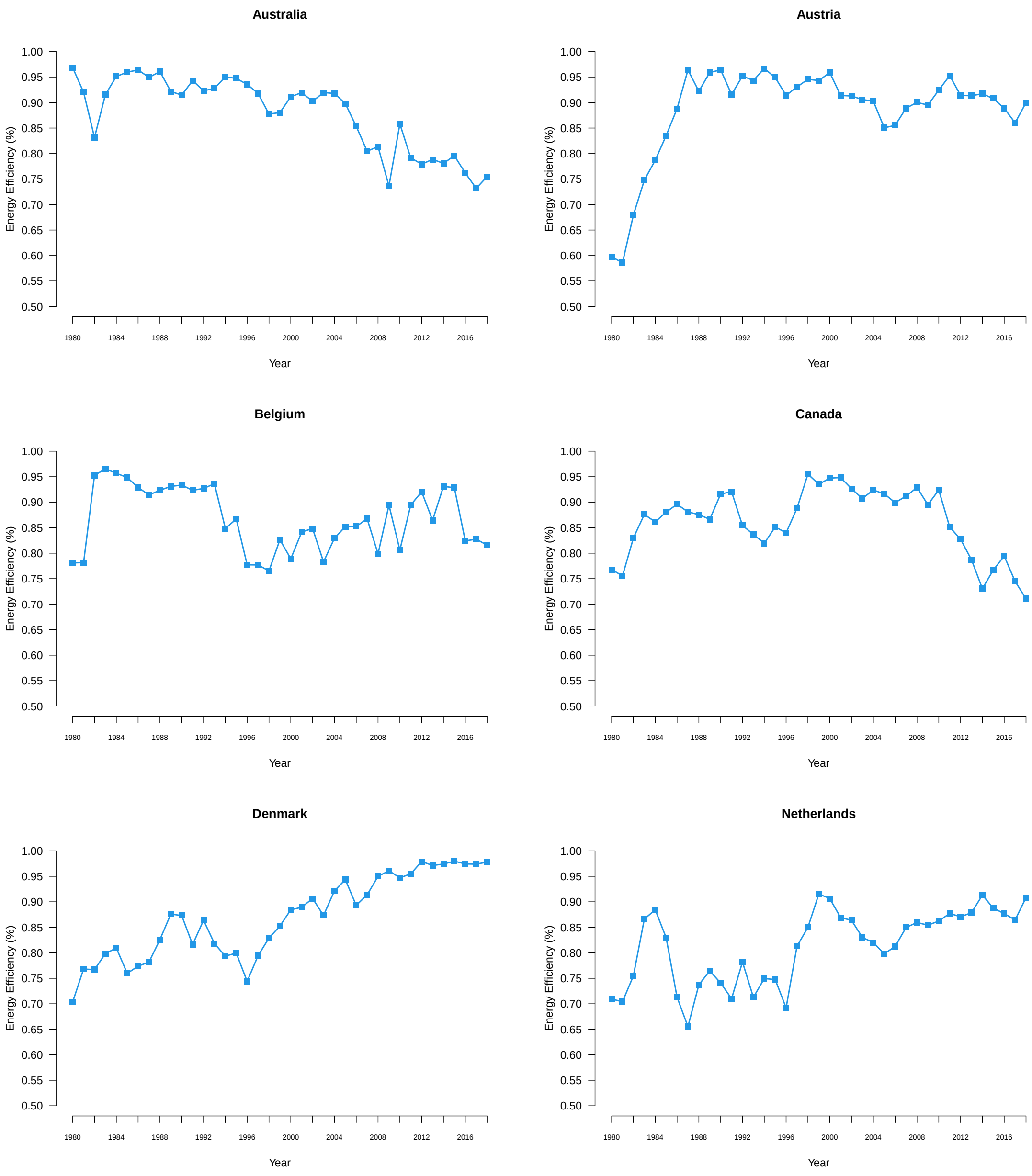


Figure 3.3: Intertemporal Energy Efficiency (1980-2018)

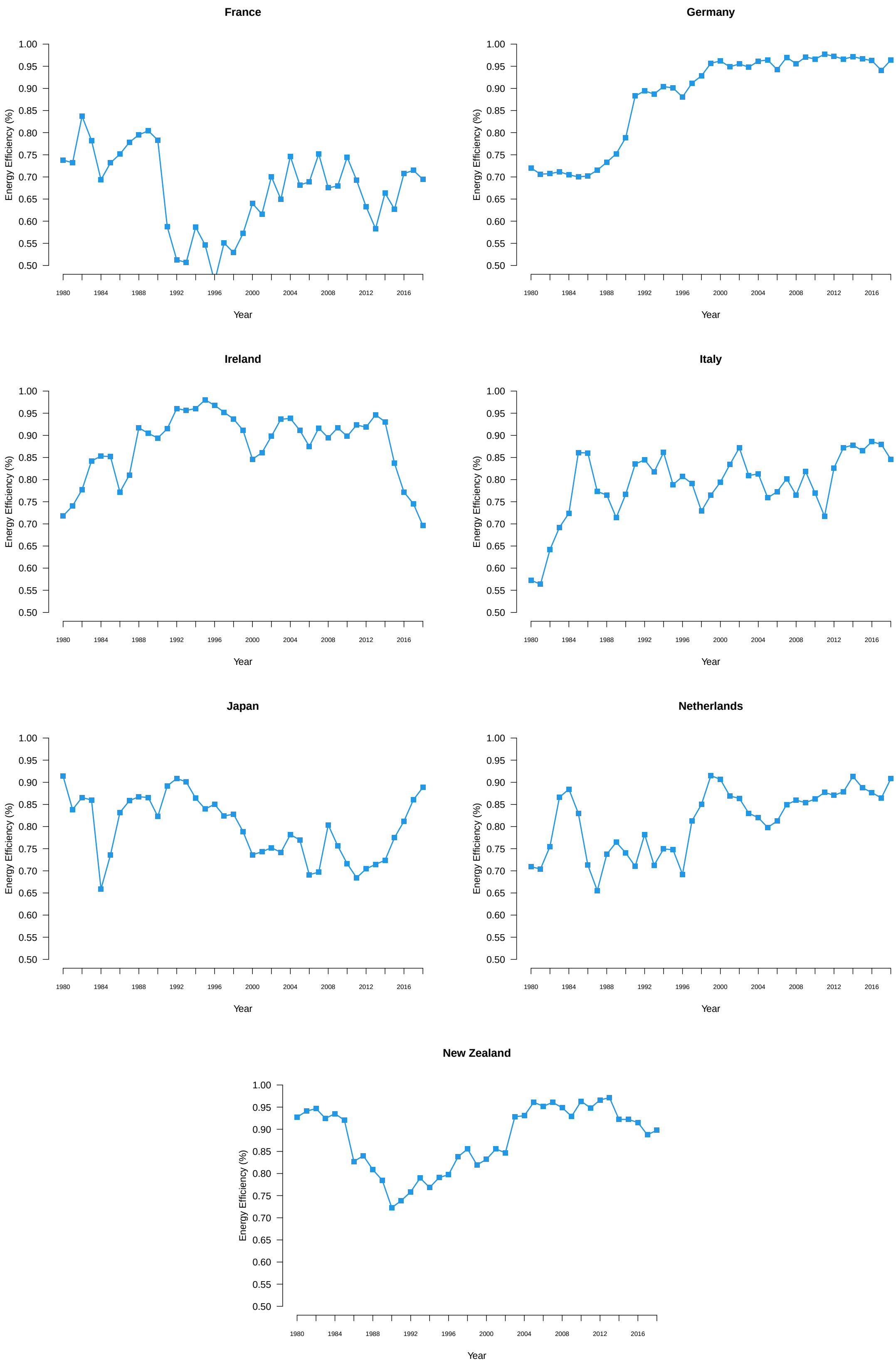


Figure 3.4: Intertemporal Energy Efficiency (1980-2018)

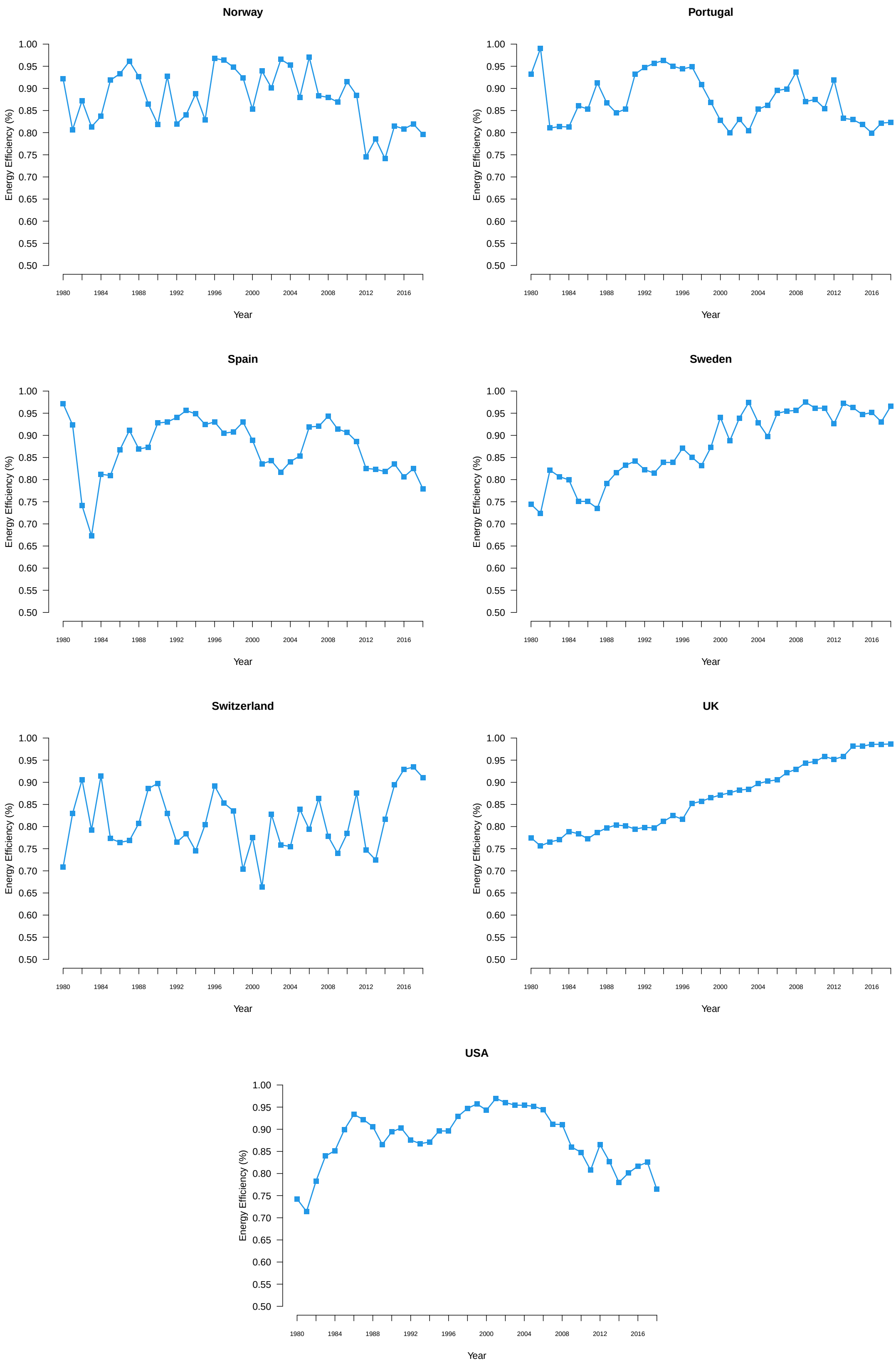


Figure 3.5: Intertemporal Energy Efficiency (1980-2018)

Bibliography

Adetutu, M. O., Glass, A. J., and Weyman-Jones, T. G. (2016). Economy-wide estimates of rebound effects: Evidence from panel data. *The Energy Journal*, 37(3).

Ahn, S. C., Good, D. H., and Sickles, R. C. (2000). Estimation of long-run inefficiency levels: a dynamic frontier approach. *Econometric Reviews*, 19(4):461–492.

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

Amsler, C., Prokhorov, A., and Schmidt, P. (2014). Using copulas to model time dependence in stochastic frontier models. *Econometric Reviews*, 33(5-6):497–522.

Amsler, C., Prokhorov, A., and Schmidt, P. (2016). Endogeneity in stochastic frontier models. *Journal of Econometrics*, 190(2):280–288.

Anderson, T. W. and Hsiao, C. (1981). Estimation of dynamic models with error components. *Journal of the American Statistical Association*, 76(375):598–606.

Anderson, T. W. and Hsiao, C. (1982). Formulation and estimation of dynamic models using panel data. *Journal of Econometrics*, 18(1):47–82.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2):277–297.

Artuc, E., Brambilla, I., and Porto, G. (2022). Patterns of labour market adjustment to trade shocks with imperfect capital mobility. *The Economic Journal*, 132(646):2048–2074.

Artuç, E., Chaudhuri, S., and McLaren, J. (2010). Trade shocks and labor adjustment: A structural empirical approach. *American economic review*, 100(3):1008–45.

Assaf, A. G. and Tsionas, M. G. (2019). Bayesian dynamic panel models for tourism research. *Tourism Management*, 75:582–594.

Ayed-Mouelhi, R. B. and Goaied, M. (2003). Efficiency measure from dynamic stochastic production frontier: Application to tunisian textile, clothing, and leather industries. *Econometric Reviews*, 22(1):93–111.

Belotti, F. and Ilardi, G. (2018). Consistent inference in fixed-effects stochastic frontier models. *Journal of Econometrics*, 202(2):161–177.

Bergeaud, A. and Ray, S. (2021). Adjustment costs and factor demand: New evidence from firms' real estate. *The Economic Journal*, 131(633):70–100.

Berner, A., Bruns, S., Moneta, A., and Stern, D. I. (2022). Do energy efficiency improvements reduce energy use? empirical evidence on the economy-wide rebound effect in europe and the united states. *Energy Economics*, 110:105939.

Bhattacharyya, A. (2012). Adjustment of inputs and measurement of technical efficiency: A dynamic panel data analysis of the egyptian manufacturing sectors. *Empirical Economics*, 42(3):863–880.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1):115–143.

Boyd, G. A. (2008). Estimating plant level energy efficiency with a stochastic frontier. *The Energy Journal*, 29(2).

Brockway, P. E., Sorrell, S., Semieniuk, G., Heun, M. K., and Court, V. (2021). Energy efficiency and economy-wide rebound effects: A review of the evidence and its implications. *Renewable and Sustainable Energy Reviews*, 141:110781.

Bruns, S. B., Moneta, A., and Stern, D. I. (2021). Estimating the economy-wide rebound effect using empirically identified structural vector autoregressions. *Energy Economics*, 97:105158.

Cave, J., Chaudhuri, K., and Kumbhakar, S. C. (2022). Dynamic firm performance and estimator choice: A comparison of dynamic panel data estimators. *European Journal of Operational Research*.

Centorrino, S. and Pérez-Urdiales, M. (2021). Maximum likelihood estimation of stochastic frontier models with endogeneity. *Journal of Econometrics*.

Chen, Y.-Y., Schmidt, P., and Wang, H.-J. (2014). Consistent estimation of the fixed effects stochastic frontier model. *Journal of Econometrics*, 181(2):65–76.

Cooper, R. W. and Haltiwanger, J. C. (2006). On the nature of capital adjustment costs. *The Review of Economic Studies*, 73(3):611–633.

Cragg, M. I. and Dyck, I. A. (1999). Management control and privatization in the united kingdom. *The RAND Journal of Economics*, pages 475–497.

Emvalomatis, G. (2012a). Adjustment and unobserved heterogeneity in dynamic stochastic frontier models. *Journal of Productivity Analysis*, 37(1):7–16.

Emvalomatis, G. (2012b). Productivity growth in german dairy farming using a flexible modelling approach. *Journal of Agricultural Economics*, 63(1):83–101.

Feng, G., Wang, C., and Serletis, A. (2018). Shadow prices of co2 emissions at us electric utilities: a random-coefficient, random-directional-vector directional output distance function approach. *Empirical Economics*, 54(1):231–258.

Fernandez, C., Osiewalski, J., and Steel, M. F. (1997). On the use of panel data in stochastic frontier models with improper priors. *Journal of Econometrics*, 79(1):169–193.

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

Filippini, M. and Hunt, L. C. (2015). Measurement of energy efficiency based on economic foundations. *Energy Economics*, 52:S5–S16.

Galán, J. E., Veiga, H., and Wiper, M. P. (2015). Dynamic effects in inefficiency: Evidence from the colombian banking sector. *European Journal of Operational Research*, 240(2):562–571.

Getachew, L. and Sickles, R. C. (2007). The policy environment and relative price efficiency of egyptian private sector manufacturing: 1987/88–1995/96. *Journal of Applied Econometrics*, 22(4):703–728.

Greene, W. (2005a). Fixed and random effects in stochastic frontier models. *Journal of Productivity Analysis*, 23(1):7–32.

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

Greene, W. H. (1990). A gamma-distributed stochastic frontier model. *Journal of Econometrics*, 46(1-2):141–163.

Greene, W. H. (2008). The econometric approach to efficiency analysis. *The measurement of productive efficiency and productivity growth*, 1(1):92–250.

Griffiths, W. E. and Hajargasht, G. (2016). Some models for stochastic frontiers with endogeneity. *Journal of Econometrics*, 190(2):341–348.

Groth, C. and Khan, H. (2010). Investment adjustment costs: An empirical assessment. *Journal of Money, Credit and Banking*, 42(8):1469–1494.

Hall, R. E. (2004). Measuring factor adjustment costs. *The Quarterly Journal of Economics*, 119(3):899–927.

Hamermesh, D. S. (1995). Labour demand and the source of adjustment costs. *The Economic Journal*, 105(430):620–634.

Hamermesh, D. S. and Pfann, G. A. (1996). Adjustment costs in factor demand. *Journal of Economic Literature*, 34(3):1264–1292.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, pages 1029–1054.

Holtz-Eakin, D., Newey, W., and Rosen, H. S. (1988). Estimating vector autoregressions with panel data. *Econometrica: Journal of the Econometric Society*, pages 1371–1395.

Hsiao, C., Pesaran, M., and Tahmiscioglu, A. (1999). Bayes estimation of short-run coefficients in dynamic panel data models. In *Analysis of Panel Data and Limited Dependent Variable Models*.

Jafari, M., Stern, D. I., and Bruns, S. B. (2022). How large is the economy-wide rebound effect in middle income countries? evidence from iran. *Ecological Economics*, 193:107325.

Jonuzaj, M. and Tsionas, M. (2023). Bayesian inference in dynamic panel stochastic frontier models. *Lancaster University Working Papers*.

Kumbhakar, S. C. and Lovell, C. K. (2000). *Stochastic Frontier Analysis*. Cambridge university press.

Kumbhakar, S. C., Parmeter, C. F., and Zelenyuk, V. (2020). Stochastic frontier analysis: Foundations and advances i. *Handbook of production economics*, pages 1–40.

Kumbhakar, S. C. and Tsionas, M. G. (2020). On the estimation of technical and allocative efficiency in a panel stochastic production frontier system model: Some new formulations and generalizations. *European Journal of Operational Research*.

Kutlu, L. (2010). Battese-coelli estimator with endogenous regressors. *Economics Letters*, 109(2):79–81.

Kutlu, L., Tran, K. C., and Tsionas, M. G. (2019). A time-varying true individual effects model with endogenous regressors. *Journal of Econometrics*, 211(2):539–559.

Lai, H.-p. and Kumbhakar, S. C. (2020). Estimation of a dynamic stochastic frontier model using likelihood-based approaches. *Journal of Applied Econometrics*, 35(2):217–247.

Liu, F., Zhang, P., Erkan, I., and Small, D. S. (2017). Bayesian inference for random coefficient dynamic panel data models. *Journal of Applied Statistics*, 44(9):1543–1559.

Lucas Jr, R. E. (1967). Adjustment costs and the theory of supply. *Journal of Political Economy*, 75(4, Part 1):321–334.

Meeusen, W. and van Den Broeck, J. (1977). Efficiency estimation from cobb-douglas production functions with composed error. *International Economic Review*, pages 435–444.

Nerlove, M. (1958). *Distributed lags and demand analysis for agricultural and other commodities*. US Government Printing Office.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, pages 1417–1426.

Nickell, S., Nicolitsas, D., and Dryden, N. (1997). What makes firms perform well? *European economic review*, 41(3-5):783–796.

Nickell, S. J. (1996). Competition and corporate performance. *Journal of Political Economy*, 104(4):724–746.

Orea, L., Llorca, M., and Filippini, M. (2015). A new approach to measuring the rebound effect associated to energy efficiency improvements: An application to the us residential energy demand. *Energy Economics*, 49:599–609.

Paul, S. and Shankar, S. (2018). On estimating efficiency effects in a stochastic frontier model. *European Journal of Operational Research*, 271(2):769–774.

Paul, S. and Shankar, S. (2019). Estimating efficiency effects in a panel data stochastic frontier model. *Journal of Productivity Analysis*, pages 1–18.

Pesaran, M. H. and Smith, R. (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of econometrics*, 68(1):79–113.

Saunders, H. D. (2000). A view from the macro side: rebound, backfire, and khazzoom–brookes. *Energy Policy*, 28(6-7):439–449.

Schmidt, P. and Sickles, R. C. (1984). Production frontiers and panel data. *Journal of Business & Economic Statistics*, 2(4):367–374.

Stern, D. I. (2012). Modeling international trends in energy efficiency. *Energy Economics*, 34(6):2200–2208.

Stern, D. I. (2020). How large is the economy-wide rebound effect? *Energy Policy*, 147:111870.

Stevenson, R. E. (1980). Likelihood functions for generalized stochastic frontier estimation. *Journal of Econometrics*, 13(1):57–66.

Tajudeen, I. A. (2021). The underlying drivers of economy-wide energy efficiency and asymmetric energy price responses. *Energy Economics*, 98:105222.

Tanner, M. A. and Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American statistical Association*, 82(398):528–540.

Theil, H. and Goldberger, A. S. (1961). On pure and mixed statistical estimation in economics. *International Economic Review*, 2(1):65–78.

Treadway, A. B. (1970). Adjustment costs and variable inputs in the theory of the competitive firm. *Journal of Economic Theory*, 2(4):329–347.

Tsionas, E. G. (2001). Posterior analysis of stochastic frontier models with truncated normal errors. *Computational Statistics*, 16(4):559–575.

Tsionas, E. G. (2002). Stochastic frontier models with random coefficients. *Journal of Applied Econometrics*, 17(2):127–147.

Tsionas, E. G. (2006). Inference in dynamic stochastic frontier models. *Journal of Applied Econometrics*, 21(5):669–676.

Tsionas, M., Patel, P. C., and Guedes, M. J. (2022). Endogenous efficiency of the dynamic profit maximization in the intertemporal production models of venture behavior. *International Journal of Production Economics*, 246:108411.

Tsionas, M. G. and Kumbhakar, S. C. (2022). Productivity and performance: A gmm approach. *Oxford Bulletin of Economics and Statistics.*

Tsionas, M. G., Malikov, E., and Kumbhakar, S. C. (2020). Endogenous dynamic efficiency in the intertemporal optimization models of firm behavior. *European Journal of Operational Research*, 284(1):313–324.

Tsionas, M. G. and Mamatzakis, E. (2019). Further results on estimating inefficiency effects in stochastic frontier models. *European Journal of Operational Research*, 275(3):1157–1164.

Tsionas, M. G. and Tzeremes, N. G. (2021). The degree of internationalization and firm productivity: Empirical evidence from large multinationals. *British Journal of Management*.

Van den Broeck, J., Koop, G., Osiewalski, J., and Steel, M. F. (1994). Stochastic frontier models: A bayesian perspective. *Journal of Econometrics*, 61(2):273–303.

Wang, H.-J. and Ho, C.-W. (2010). Estimating fixed-effect panel stochastic frontier models by model transformation. *Journal of Econometrics*, 157(2):286–296.

Wang, H.-J. and Schmidt, P. (2002). One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels. *Journal of Productivity Analysis*, 18(2):129–144.