Bayesian Inference in Dynamic Panel Stochastic Frontier Models

Mariol Jonuzaj[*] Mike Tsionas[†] Marwan Izzeldin[‡]

January 2026

Abstract

The paper develops a dynamic panel stochastic frontier model that incorporates firms' intertemporal decision behaviour and short-run stagnant adjustments to the production process. Its dynamic specification recognises short-run output adjustment costs, where final output may be only partially adjusted to the optimum level. In nesting previous panel stochastic frontier models, our new approach delivers a flexible framework that accommodates heterogeneous technologies and latent time-varying inefficiency effects. In addition, our model handles endogeneity issues related to flexible inputs. Model inference is based on a Bayesian framework, where Markov Chain Monte Carlo (MCMC) techniques are utilized. Through extensive simulations, we demonstrate the robustness of the model in small and moderate samples. Last, we present our model in an empirical example, analysing publicly listed UK companies operating in the manufacturing and construction sector over the period 2004-2022. A general finding is that most firms exhibit stagnant production processes, with the half-life for adjusting supply to be as high as 6 quarters. The estimated average technical efficiency is 89%. Our findings underscore the importance of accounting for dynamic frictions and heterogeneity when evaluating firm performance and designing productivity-enhancing policies.

Keywords: Dynamic Panel Data, Efficiency Analysis, MCMC, Partial Adjustment, Stochastic Frontier Models

JEL Classifications: C23, C50, D24

[*]Department of Economics, Royal Holloway, University of London, TW20 0EX, UK, Email address: mariol.jonuzaj@rhul.ac.uk

[†]Montpellier Business School, Université de Montpellier, France and Department of Economics, Lancaster University Management School, LA1 4YX, UK, Email address: m.tsionas@lancaster.ac.uk

[‡]Department of Economics, Lancaster University Management School, LA1 4YX, UK, Email address: m.izzeldin@lancaster.ac.uk

1 Introduction

Measuring firm productivity and technical efficiency is vital for firms seeking to assess production performance, identify cost-saving potentials, and enhance competitiveness. Stochastic frontier analysis provides a standard econometric framework 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 statistical and econometric analysis to consider potential technical inefficiencies within the production process, which can arise from managerial errors and general operational practices.

Although stochastic frontier models[1] have been extended in various directions, the dominant empirical frontier literature remains largely static, implicitly assuming that firm can adjust output instantaneously and without costs. In practice, firms are often operate in regulated and rapidly changing markets, where production adjustment is neither costless nor instantaneous. Adjustment costs may arise from quasi-fixed inputs and organisational frictions, such as the time required to train workers, install and learn new technologies, or reorganise production in response to regulatory, trade, macroeconomic, and labour market conditions (Hall (2004), Cooper and Haltiwanger (2006), Groth and Khan (2010), Bergeaud and Ray (2021), Artuç et al. (2010), Artuc et al. (2022)).

As an example, in the manufacturing and construction sectors production processes are often affected by market and environmental regulations that require firms to adjust their operating practices. In particular, many energy-intensive firms in the EU and the UK are subject to the Emissions Trading System, which requires gradual changes in production methods toward more environmentally friendly practices. These regulatory changes take time to implement and can temporarily limit output, even when firms operate efficiently.

It is apparent that in any production process, even in the absence of productive inefficiencies, the inherently sluggish adjustment associated with changes in production practices and operating conditions can directly affect the production process and, consequently, the final produced output. As a result, econometric specifications that ignore such production dynamics can generate misleading inferences regarding the economic performance of a production unit, by conflating adjustment frictions with technical inefficiency.

In this line, this paper fills the gap in the literature by introducing a reduced-form econometric model that explicitly accounts for these frictions. Specifically, we present a generalised panel stochastic frontier model that distinguishes between a latent long-run production frontier and a short-run attainable frontier, the latter reflecting sluggish output adjustment arising from

[1] For an excellent recent literature review, see the textbook by Kumbhakar et al. (2020).

changes in business environments. To achieve this, we employ a Nerlove-type partial adjustment mechanism in which observed output follows a gradual intertemporal path towards the long-run equilibrium level. The proposed specification delivers a dynamic panel model with time-varying inefficiency effects, while allowing inefficiency to be measured conditional on the presence of output adjustment frictions. Furthermore, the model addresses input endogeneity, as decisions regarding flexible inputs may be correlated with production shocks.

Specifically, we embed the dynamic panel frontier model within a random coefficient setting, so that firms exhibit heterogeneous adjustment speeds and production input elasticities. The aim is to introduce a flexible specification in which firms differ both in their production technologies and in their speed of adjustment towards the long-run equilibrium. In addition, we allow latent time-varying technical inefficiencies to be present in the model. Technical inefficiency is modeled using distributional assumptions, enabling the estimation of efficiency scores relative to the firm's own production frontier.

We propose a Bayesian hierarchical structure which can effectively account for the various heterogeneous effects on the data and provide statistical evidence for all the parameters and the inefficiency scores. Explicitly, model inference and estimation rely on a Bayesian framework, utilizing Markov Chain Monte Carlo (MCMC) techniques with data augmentation. We suggest proper prior elicitation and derive all the corresponding conditional distributions. We adopt a Gibbs Sampler iteration approach where the marginal posterior distributions of each parameter can be effectively obtained. Using simulated data, we test the performance of the proposed Gibbs Sampler under various settings, and we illustrate that our Bayesian technique performs very well in both small and moderate samples.

Finally, we apply the proposed model to an empirical study where we analyse publicly listed UK companies operating in the manufacturing and construction sector covering the period 2004-2022. The timeframe of our study holds significant practical importance due to several major economic events, including the 2007-2009 financial crisis, the UK green and digital transition, the post Brexit referendum period and the COVID pandemic. Consequently, our study provides important empirical insights regarding the firms' performance and the adjustment process that has taken place during this transition period. To begin with, our econometric analysis reveals considerable heterogeneous effects among the firms, with the majority of them exhibiting significant stagnant production adjustments. Specifically, the average "half-life" adjustment period is 5 months, and it can be up to 6 quarters. The estimated average short-run return to scales is approximately 0.824, indicating that the short-run average curves are upward-sloping. This reveals that firms during the period of our study undertook considerable adjustments, which introduced financial pressure during the first stage of transition. Additionally, in terms of technical efficiencies, the average efficiency score is 89%, indicating that even though firms produce near their frontier, they still can decrease their average cost by 12%, ceteris paribus. Regarding policy implications, the results reveal that during periods of economic turbulence and transition to more digitalized and carbon-neutral environments, policymakers should take into account the stagnant behaviour of production and propose flexible and economically sustainable policies. Tight economic policies and regulations can harm firms and eventually affect consumers through the price channel.

The rest of the paper is organized as follows. In section 2, we present the relevance and contribution of our study. In section 3, the conceptualization framework is presented. In section 4, we illustrate the proposed model. In section 5, we outline the proposed Bayesian approach. In section 6, we test the performance of model using Monte Carlo simulations. In section 7, we present the empirical application. Last, in section 8, the conclusion is stated.

2 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, the model is a generalized extension of panel stochastic frontier models that have been utilized in empirical studies. Specifically, our model introduces production dynamics in a manner that nests static panel models presented in Tsionas (2002), Greene (2005a,b), Wang and Ho (2010), Chen et al. (2014), Belotti and Ilardi (2018), Kutlu et al. (2019), among others. All the aforementioned papers introduce static specifications where both heterogeneity and inefficiency are treated in various forms.

Moreover, the proposed model extends dynamic panel production models and allows time-varying inefficiency effects to be present. For instance, Nickell (1996), Nickell et al. (1997) and Ayed-Mouelhi and Goaied (2003) are some early studies that utilize dynamic panel production functions to incorporate the short-run adjustment frictions and examine the firms' total factor productivity. All these studies do not allow for heterogeneous production functions and ignore the time-varying nature of technical inefficiency. By contrast, our model permits flexible specification associated with different technological possibilities; and allows time-varying technical inefficiency. In addition, we separate the unobserved heterogeneity effects from the inefficiency effects. This is quite important in practice since failing to do so, will result in distorted 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 (2012, 2020), Amsler et al. (2014), Galán et al. (2015), Lai and Kumbhakar (2020), Tsionas and Kumbhakar (2023). Our conceptualization is similar to this literature, but we allow sluggish production adjustments not only to affect the inefficiency path but also the whole production process. Specifically, our proposed autoregressive model permits input and inefficiency changes to have an intertemporal impact on production output.

Last, the proposed model addresses input endogeneity problems related to the timing of each input decision. In the production economics literature (see Ackerberg et al. (2015), Gandhi et al. (2020) and the bibliography cited therein), it is well understood that flexible inputs (e.g. intermediate expenditure) are affected instantaneously by production shocks. Thus, failing to account for this could lead to inconsistent estimates. To address this issue, our proposed model follows recent econometric advancements introduced by Kutlu (2010), Kim and Kim (2011), Amsler et al. (2016), Tsionas (2017), Kutlu et al. (2019), Tsionas et al. (2023), among others.

3 Conceptualization Framework

To illustrate the idea of dynamic production frontiers, we assume a sector where N firms are operating over a time period T. Each firm uses 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 that reflects the firm's targeted level of output. This production process can be described as:

$$Y_{it}^{*} = f(\boldsymbol{X}_{it}; \boldsymbol{\beta}^{*}) \exp(v_{it}^{*}) \tag{1}$$

where $f(.)$ can be any production function such as a Cobb-Douglas, $\boldsymbol{X}_{it}$ is the set of inputs (e.g. labour, capital, intermediate inputs) used by firm i at time t, $\boldsymbol{\beta}^{*}$ is the vector of the technological parameters and $\exp(v_{it}^{*})$ denotes the output-oriented production shocks. Here, we note that in the long-run specification, all firms are assumed to be fully efficient[2].

In logarithm form, the long-run production frontier will be:

$$y_{it}^{*} = a_{i}^{*} + \boldsymbol{x}_{it}' \boldsymbol{\beta}^{*} + v_{it}^{*} \tag{2}$$

[2] A parallel strand of the literature, allows for the presence of long-run inefficiency effects. Models incorporating persistent inefficiency can be found in Tsionas and Kumbhakar (2014), Filippini and Greene (2016) and the related literature cited there.

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 random productivity shocks.

As discussed above, firms operating under regulatory constraints, or organisational frictions are generally unable to adjust their observed output instantaneously to the long-run target level y_{it}^*. Even when firms operate efficiently, short-run production may fall short of the long-run frontier due to adjustment frictions. To capture this distinction, we introduce a short-run attainable level of log-output, denoted by $y_{it}^\dagger$, which reflects the maximum output that can be achieved in the short-run given existing adjustment constraints but in the absence of technical inefficiency. Actual log-output y_{it}, may further deviate from the short-run frontier due to technical inefficiency, $u_{it} \geq 0$, which measures the shortfall of realised output from the short-run attainable level. Formally:

$$y_{it} = y_{it}^\dagger - u_{it}$$

We model the evolution of short-run attainable output using a partial adjustment mechanism à la Nerlove (1958), given by:

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

or equivalently:

$$y_{it} = \lambda y_{it}^* + (1 - \lambda) y_{it-1} - u_{it} \tag{3}$$

where y_{it} is the actual log-output at time t, y_{it-1} is the log-output produced at the lagged period $t-1$, u_{it} is the technical inefficiency, y_{it}^* is the desired or targeted log-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 when $\lambda = 1$, $y_{it}^\dagger$ adjusts instantaneously to the long-run level y_{it}^*, and the only source of output fall are the usual inefficiency effects. On the other hand, when $0 < \lambda < 1$, the process is subject to some level of inertia, causing y_{it} to deviate from y_{it}^* due to (i) the inherent adjustment frictions and (ii) the technical inefficiency effects.

Here, we should highlight, that the proposed partial adjustment mechanism does not alter the usual definition of production theory, which links current inputs to current output, but instead provides a framework for explaining and decomposing the dynamic evolution of the output gap[3]. Specifically, given the usual definition of output gap, denoted as $g_{it} \equiv y_{it}^* - y_{it}$,

[3] We would like to thank an anonymous referee for noting this concern.

we can show the following theoretical formulation:

$$
\begin{aligned}
g_{it} &\equiv y_{it}^{*} - y_{it} \\
&= y_{it}^{*} - \lambda y_{it}^{*} - (1-\lambda) y_{it-1} + u_{it} \\
&= (1-\lambda) y_{it}^{*} - (1-\lambda)(y_{it-1}^{*} - g_{it-1}) + u_{it} \\
&= (1-\lambda) g_{it-1} + (1-\lambda)(y_{it}^{*} - y_{it-1}^{*}) + u_{it}
\end{aligned}
$$

It is evident that, even in the presence of technical inefficiencies, the partial adjustment mechanism delivers a law of motion for the output gap consisting of three components:

1. the output gap correction component, where firms close part of their past output gap through learning
2. the newly generated output gap, arising from the frontier expansions and the firm's inability to adjust instantaneously to the newly available technology.
3. the inefficiency component, which reflects deviations arising from managerial practices or general internal production dis-utilities.

Last, combining the long-run equilibrium relationship in equation (2) with the short-run dynamics from equation (3), we obtain the final form of the model:

$$y_{it} = \rho y_{it-1} + \boldsymbol{x}_{it}' \boldsymbol{\beta} + a_i + v_{it} - u_{it} \tag{4}$$

where $\rho = (1-\lambda)$, $\boldsymbol{\beta} = \lambda \boldsymbol{\beta}^{*}$, $a_i = \lambda a_i^{*}$ and $v_{it} = \lambda v_{it}^{*}$. In this form, it is clear that the partial adjustment mechanism delivers a dynamic panel data stochastic frontier model. Under this setting, the proposed model permits past input and inefficiency shocks to have an intertemporal effect on the production process. This is quite important in practice, as it allows empirical researchers to distinguish the short-run and the long-run production dynamics. For instance, the magnitude of the autoregressive parameter ρ serves as an indicator of how fast the production process adjusts to the desired level over-time. This dynamic specification allows the identification of the short-run returns to scale, which is important for revealing the additional costs that firms encounter when adapting to new production requirements. Moreover, the model allows for the estimation of efficiency scores, taking into account the fact that part of the output drop should be attributed to the stagnant nature of the production process rather than to firm production inefficiencies.

Estimation and inference in dynamic panel data models typically involves instrumental variable (IV) and Generalized Methods of Moments (GMM) techniques. 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 System-GMM techniques[4]. While the previously mentioned approaches may offer a consistent means for parameter estimation, it is important to note that obtaining point estimate for the inefficiency effect could present challenges and complexities. Briefly, GMM methods involve first difference transformations which could raise identification problems as the probability density functions of inefficiency change Δu_{it} usually has an unknown function form (see Belotti and Ilardi (2018)). For this reason, we opt for likelihood and Bayesian based models and estimate the model utilising the in-levels form.

4 A Generalisation of the Model

To generalise the model in (4), one way to capture unobserved heterogeneity in the production process is to introduce a random coefficients stochastic frontier model where the heterogeneity 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. This is broadly recognized in the empirical literature, as factor input elasticities can be quite heterogeneous among the firms reflecting the heterogeneous nature of marginal revenues and the marginal costs. For this reason, we opt for a heterogeneous coefficient setting to allow flexibility in our empirical model. 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.

In addition, we allow for the presence of flexible production inputs, which indicates that a firm's demand for these inputs will be affected by random production shocks instantaneously. This specification is in line with the recent on production economics that has been described at Olley and Pakes (1996), Levinsohn and Petrin (2003), Ackerberg et al. (2015), Gandhi et al. (2020), among many others.

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

$$y_{it} = \rho_i y_{it-1} + \boldsymbol{x}'_{it}\boldsymbol{\beta}_i + q_{it}\gamma_i + a_i + v_{it} - u_{it} \tag{5}$$

where q_{it} is the log-value of the flexible input variable. Here, we note that for simplicity and clarity, we allow for a single flexible input, as it is common in production models to consider

[4]Recently, Cave et al. (2022) illustrated an extensive Monte Carlo simulation study, where the statistical performance of different dynamic panel estimators is evaluated.

only the intermediate input as flexible[5].

Thus, throughout the paper, we make the following model assumptions:

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

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 5 *The input* q_{it} *is correlated with the error term* v_{is} *for* $t = s$*, but independent of* v_{is} *for* $t \neq s$.

Assumption 1 is common in the stochastic frontier literature, allowing production shocks to have a symmetric distribution and not being serially correlated. **Assumption 2** derives naturally from the fact that the inefficiency term cannot take negative values. In addition, **Assumption 3** is dominant in the productivity and stochastic frontier literature and is vital for the derivation of the marginal distribution of the composite term $\varepsilon_{it} = v_{it} - u_{it}$. **Assumption 4** implies that the input choice is uncorrelated with random production and inefficiency shocks[6]. Last, **Assumption 5** describes the endogeneity issue related to the flexible input. As described above, only instantaneous production shocks affect the flexible input decision. In general, we argue that all the above assumptions can be considered minimalistic to identify the parameters of interest and the latent inefficiency scores. Moreover, all model assumptions are aligned with the usual profit maximization behaviour.

Following the literature of stochastic frontier models with endogenous inputs, we address the endogeneity of q_{it} by coupling the model in (5), with the following reduced form equation:

$$q_{it} = \boldsymbol{z}_{it}' \boldsymbol{\lambda}_i + e_{it} \tag{6}$$

where $\boldsymbol{z}_{it}$ is a vector of instruments and e_{it} is the residual term. In practice, equation (6) relates the endogenous input q_{it}, to a vector of exogenous variables $\boldsymbol{z}_{it}$ that are orthogonal to

[5] Of course, the model can include multiple flexible inputs.

[6] In most empirical applications, researchers control for time and firm-specific effects which alleviate omitted variable issues. In addition, we could further relax the strict exogeneity assumption and allow the input vector $\boldsymbol{x}_{it}$ to be correlated only with past production shocks. However, given that v_{it} is not serially correlated, this will not affect our estimates.

production shocks v_{it}. Specifically, the vector $\boldsymbol{z}_{it}$ may include relevant variables such as input and raw material prices, transportation prices and others[7], as these factors directly influence the flexible input q_{it}. Additionally, these variables are orthogonal to the firm specific production shocks, as prices are formulated by general market dynamics and firms behave as price takers.

Moving forward, it is important to allow equations (5) and (6) to be determined jointly in a system of equations framework by introducing dependency between v_{it} and e_{it}. We do that by following Kutlu (2010), Kim and Kim (2011), Amsler et al. (2016), Tsionas (2017), Kutlu et al. (2019), Tsionas et al. (2023) among others, and we introduce the following structure:

$$\begin{bmatrix} v_{it} \\ e_{it} \end{bmatrix} \sim \boldsymbol{\mathcal{N}} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_v^2 & \theta\sigma_v\sigma_e \\ \theta\sigma_v\sigma_e & \sigma_e^2 \end{bmatrix} \right)$$

where θ is the correlation between v_{it} and e_{it}.

Using Cholesky decomposition on the variance-covariance matrix of $[v_{it}, e_{it}]'$ we can obtain the final form of the frontier equation as:

$$y_{it} = \rho_i y_{it-1} + \boldsymbol{x}'_{it}\boldsymbol{\beta}_i + q_{it}\gamma_i + a_i + \theta\frac{\sigma_v}{\sigma_e}(q_{it} - \boldsymbol{z}'_{it}\boldsymbol{\lambda}_i) + \sigma_v\sqrt{1-\theta^2}\omega_{it} - u_{it} \tag{7}$$

where $\omega_{it} \sim \mathcal{N}(0,1)$. It is important to notice that the frontier equation in (7) delivers a specification where all variables are independent of the error term ω_{it}. In addition, the term $\theta\frac{\sigma_v}{\sigma_e}(q_{it} - \boldsymbol{z}'_{it}\boldsymbol{\lambda}_i)$ works as a bias correction term.

Thus, we can write the full model in matrix form as:

$$\begin{gathered} \boldsymbol{y}_i = \rho_i \boldsymbol{y}_{i,-1} + \boldsymbol{X}_i\boldsymbol{\beta}_i + \boldsymbol{q}_i\gamma_i + \theta\frac{\sigma_v}{\sigma_e}(\boldsymbol{q}_i - \boldsymbol{Z}_i\boldsymbol{\lambda}_i) + \boldsymbol{\omega}_i - \boldsymbol{u}_i \\ \boldsymbol{q}_i = \boldsymbol{Z}_i\boldsymbol{\lambda}_i + \boldsymbol{e}_i \end{gathered} \tag{8}$$

with

$$\boldsymbol{\omega}_i \sim \mathcal{N}(\boldsymbol{0}, \sigma_v^2(1-\theta^2)\boldsymbol{I}) \ , \ \boldsymbol{e}_i \sim \mathcal{N}(\boldsymbol{0}, \sigma_e^2\boldsymbol{I})$$

for each $i = 1, 2, \ldots, N$ where $\boldsymbol{y}_i = [y_{i1}, y_{i2}, \ldots, y_{iT}]'$ is the $T \times 1$ matrix of the dependent variable, $\boldsymbol{y}_{i,-1} = [y_{i0}, y_{i1}, \ldots, y_{iT-1}]'$ is the $T \times 1$ matrix of the lagged log-output, $\boldsymbol{X}_i = [\boldsymbol{1}', \boldsymbol{x}'_{i1}, \boldsymbol{x}'_{i2}, \ldots, \boldsymbol{x}'_{iT}]'$ is a $T \times K_1$ matrix of the log-inputs and the constant, $\boldsymbol{q}_i = [q_{i1}, q_{i2}, \ldots, q_{iT}]$ is the $T \times 1$ matrix of the endogenous log-input and $\boldsymbol{Z}_i = [\boldsymbol{1}', \boldsymbol{z}'_{i1}, \boldsymbol{z}'_{i2}, \ldots, \boldsymbol{z}'_{iT}]'$ is the $T \times K_2$ matrix containing a constant and the instrumental variables. In addition, $\boldsymbol{\omega}_i = [\omega_{i1}, \omega_{i2}, \ldots, \omega_{iT}]$ is the $T \times 1$ matrix of the symmetric error term, $\boldsymbol{e}_i = [e_{i1}, e_{i2}, \ldots, e_{iT}]'$ is the $T \times 1$ matrix of error term of the second equation, $\boldsymbol{I}$ is a $T \times T$ identity matrix, and $\boldsymbol{u}_i = [u_{i1}, u_{i2}, \ldots, u_{iT}]'$ is the $T \times 1$ matrix of the technical inefficiency term. Regarding the initial values conditions y_{i0} for each $i = 1, 2, \ldots, N$, we depart from the approach used in Hsiao et al. (1999), where the

[7]Other potential valid instruments include various supply chain indices published by various central banks.

initial values are considered as fixed and known, and instead, we are treating them as latent variables to be estimated.

Last, to complete the model, for the inefficiency term, we make use of the half-normal distribution, as proposed by Aigner et al. (1977)[8], with:

$$\boldsymbol{u}_i \sim \boldsymbol{\mathcal{N}}^{+}(\boldsymbol{0}, \sigma_u^2 \boldsymbol{I}) \tag{9}$$

for each $i = 1, 2, \ldots, N$. In this study, we choose the half-normal distribution because it is a common distribution for modeling inefficiencies, and we want our modeling framework to reflect that. This approach also allows us to nest previously proposed panel stochastic frontier models. Additionally, our goal is to provide conditional posterior distributions in a known form and avoid complex numerical estimation procedures. Lastly, our choice of the half-normal distribution is driven by our empirical study. Specifically, as we examine large publicly traded firms, we aim to assign apriori some probability that firms can be highly efficient. This reflects the fact that the UK market is highly competitive and therefore large companies must be highly efficient. This feature can be achieved using the half-normal distribution, as it places some density on values close to zero.

The proposed model in (8), serves as a generalization of previous panel data stochastic frontier models that have been introduced in the literature, so far. Equation (8) clearly extends existing models by incorporating production dynamics while controlling for the presence of flexible inputs. Next, we discuss in details, how the proposed model nests existing panel frontier models.

First, we note that the for the case where $\theta = 0$, we arrive to a dynamic panel stochastic frontier model where all inputs are uncorrelated with current production shocks (predetermined). This model, remains novel as it generalizes static models under input exogeneity. Secondly, for the case where firms adjust perfectly their production to their targeted levels, viz. $\rho_i = 0$ for all $i = 1, 2, \ldots, N$ but we still retain the assumption of flexible inputs, viz. $\theta \neq 0$, our model arrives to specifications of static random coefficient frontier with input endogeneity. Similar specifications can be find at Kutlu (2010), Amsler et al. (2016), Tsionas (2017), Kutlu et al. (2019), Tsionas et al. (2023), among others. Of course, each study has specified different either the production frontier or the inefficiency term. Third, when $\rho_i = 0$ for all $i = 1, 2, \ldots, N$ and $\theta = 0$, 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

[8] 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).

firms may face different technological capabilities and failing to properly control for that could inflate the estimated inefficiency scores. In addition, when the random coefficients assumption is constrained to a fixed coefficient specification, viz $\boldsymbol{\beta}_i = \bar{\boldsymbol{\beta}}$ for all $i = 1, 2, \ldots, N$, 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)[9]. 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 $a_i = \bar{a}$ for all $i = 1, 2, \ldots, N$, 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) [10].

5 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 (8), as:

$$\begin{aligned} \boldsymbol{y}_i &= \boldsymbol{W}_i \boldsymbol{\delta}_i + \theta \frac{\sigma_v}{\sigma_e}(\boldsymbol{q}_i - \boldsymbol{Z}_i \boldsymbol{\lambda}_i) + \boldsymbol{\omega}_i - \boldsymbol{u}_i \\ \boldsymbol{q}_i &= \boldsymbol{Z}_i \boldsymbol{\lambda}_i + \boldsymbol{e}_i \end{aligned} \tag{10}$$

with

$$\boldsymbol{\omega}_i \sim \mathcal{N}(\mathbf{0}, \sigma_v^2(1-\theta^2)\boldsymbol{I}) \ , \ \boldsymbol{e}_i \sim \mathcal{N}(\mathbf{0}, \sigma_e^2 \boldsymbol{I})$$

where $\boldsymbol{W}_i = [\boldsymbol{y}_{i,-1}, \boldsymbol{X}_i, \boldsymbol{q}_i]$ will be a $T \times J$ matrix and $\boldsymbol{\delta}_i = [\rho_i, \boldsymbol{\beta}_i, \gamma_i]'$ is the $J \times 1$ matrix of unknown parameters, where $J = K_1 + 2$. Equation (10) consists of the first stage of our hierarchical structure.

5.1 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

[9] 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.

[10] The paper 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 $\mu = 0$.

distribution, with:

$$\boldsymbol{\delta_i} = \begin{bmatrix} \rho_i \\ \boldsymbol{\beta_i} \\ \gamma_i \end{bmatrix} \sim \boldsymbol{\mathcal{N}}(\bar{\boldsymbol{\delta}}, \boldsymbol{\Omega}) \tag{11}$$

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:

$$\bar{\boldsymbol{\delta}} \sim \boldsymbol{\mathcal{N}}(\boldsymbol{\mu_0}, \boldsymbol{\Lambda_0}) \tag{12}$$

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 their prior beliefs. From (12), it is obvious that one can assume a fully non-informative flat prior by imposing $\boldsymbol{\mu_0} = \mathbf{0}_{J\times 1}$ and $\boldsymbol{\Lambda_0} = 10^3 \times \boldsymbol{I}_{J\times J}$, where $\mathbf{0}_{J\times 1}$ is the $(J \times 1)$ null matrix and $\boldsymbol{I}_{J\times J}$ is a $J \times J$ identity matrix. Given that the parameter vector $\bar{\boldsymbol{\delta}}$ captures production dynamics and input elasticities, the proposed prior does not dominate the likelihood.

We, follow similar strategy for the parameter vector $\boldsymbol{\lambda}_i$ which captures the coefficient of the instrumental variables. We specify:

$$\boldsymbol{\lambda_i} \sim \boldsymbol{\mathcal{N}}(\bar{\boldsymbol{\lambda}}, \boldsymbol{A}) \tag{13}$$

where

$$\bar{\boldsymbol{\lambda}} \sim \boldsymbol{\mathcal{N}}(\boldsymbol{\lambda}_0, \boldsymbol{\Gamma}_0) \tag{14}$$

Again we impose $\boldsymbol{\lambda}_0 = \mathbf{0}_{J_2\times 1}$ and $\boldsymbol{\Gamma}_0 = 10^3 \times \boldsymbol{I}_{J_2\times J_2}$, where J_2 is the number of columns of $\boldsymbol{Z}_i$ matrix.

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

$$\boldsymbol{\Omega}|\boldsymbol{\Psi}_0, v_0 \sim \boldsymbol{\mathcal{IW}}(\boldsymbol{\Psi}_0, v_0) \tag{15}$$

with probability density function:

$$\pi(\boldsymbol{\Omega}|\boldsymbol{\Psi}_0, v_0) \propto det(\boldsymbol{\Omega})^{-\frac{v_0+J+1}{2}} exp\left\{-\frac{1}{2}tr(\boldsymbol{\Psi}_0\boldsymbol{\Omega}^{-1})\right\} \tag{16}$$

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:

$$\pi(\boldsymbol{\Omega}|\boldsymbol{\Psi}_0 = 10^{-6} \times \boldsymbol{I}_{J\times J}, v_0 = 0) \propto det(\boldsymbol{\Omega})^{-\frac{J+1}{2}} \tag{17}$$

Similarly, for the variance-covariance matrix $\boldsymbol{A}$, we use $\boldsymbol{A}|\boldsymbol{\Phi}_0, v_0 \sim \boldsymbol{\mathcal{IW}}(\boldsymbol{\Phi}_0, v_0)$. We use the non-informative Jeffrey's prior, by setting the following:

$$\pi(\boldsymbol{A}|\boldsymbol{\Phi}_0 = 10^{-6} \times \boldsymbol{I}_{J_2\times J_2}, v_0 = 0) \propto det(\boldsymbol{A})^{-\frac{J_2+1}{2}} \tag{18}$$

For the variances σ_v^2, σ_u^2 and σ_e^2, we use the Inverse-Gamma prior distribution, of the form:

$$\sigma_x^2|a_0, a_1 \sim \mathcal{IG}(a_0, a_1) \quad , \quad x \in \{v, u, e\} \tag{19}$$

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 $\pi(\sigma_x^2) \propto \frac{1}{\sigma_x^2}$. 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 $a_0 = a_1 = 10^{-2}$ (see Fernandez et al. (1997), page 186). In this way, we generate diffuse priors, widely used to indicate minimal prior knowledge and avoid bias in posterior estimates.

Regarding the correlation coefficient θ, we use a Beta prior distribution, with:

$$\theta|b_0, b_1 \sim Beta(b_0, b_1) \tag{20}$$

We generate a non-informative prior by setting $b_0 = b_1 = 1$. This leads to a non-informative uniform distribution on the support $[0, 1]$, expressing complete uncertainty about the value of parameter θ . Other potential options could be $b_0 = b_1 = 2$.

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:

$$y_{i0} \sim \mathcal{N}(\bar{y}_0, \sigma_0^2) \tag{21}$$

This states that the latent initial values conditions are distributed around a common mean $\bar{y}_0$ and a variance σ_0^2. This "knowing little" prior specification depicts the firms' initial endowments and these effects gradually vanish over time.

5.2 Posterior Analysis

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

$$p(\boldsymbol{\theta}, \{\boldsymbol{\delta}_i\}_{i\in[1,N]}, \{\boldsymbol{\lambda}_i\}_{i\in[1,N]}, \boldsymbol{u}, \boldsymbol{y}_0 | \boldsymbol{Y}, \boldsymbol{W}, \boldsymbol{Z}) \propto \left[\prod_{i=1}^{N} f(\boldsymbol{Y}_i, \boldsymbol{W}_i, \boldsymbol{Z}_i | \boldsymbol{\theta}, \boldsymbol{\delta}_i, \boldsymbol{\lambda}_i, \boldsymbol{u}_i, y_{i0}) f(\boldsymbol{\delta}_i, \boldsymbol{\lambda}_i, \boldsymbol{u}_i, y_{i0} | \boldsymbol{\theta})\right] \pi(\boldsymbol{\theta}) \tag{22}$$

where $\boldsymbol{Y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_N]'$, $\boldsymbol{W} = [\boldsymbol{W}_1, \boldsymbol{W}_2, \ldots, \boldsymbol{W}_N]'$, $\boldsymbol{Z} = [\boldsymbol{Z}_1, \boldsymbol{Z}_2, \ldots, \boldsymbol{Z}_N]'$, $\boldsymbol{u} = [\boldsymbol{u}_1, \boldsymbol{Z}_u, \ldots, \boldsymbol{u}_N]'$ and $\boldsymbol{y_0} = [y_{10}, y_{20}, \ldots, y_{N0}]'$. $p(\boldsymbol{\theta}, \{\boldsymbol{\delta}_i\}_{i\in[1,N]}, \{\boldsymbol{\lambda}_i\}_{i\in[1,N]}, \boldsymbol{u}, \boldsymbol{y}_0 | \boldsymbol{Y}, \boldsymbol{W}, \boldsymbol{Z})$ denotes the augmented posterior distribution of the structural parameters $\boldsymbol{\theta} = [\bar{\boldsymbol{\delta}}, \bar{\boldsymbol{\lambda}}, \sigma_v^2, \sigma_u^2, \sigma_e^2, \theta, \boldsymbol{\Omega}, \boldsymbol{A}]'$ and the latent parameters $\boldsymbol{\delta}_i, \boldsymbol{\lambda}_i, \boldsymbol{u}, \boldsymbol{y}_0$. $f(\boldsymbol{Y}_i, \boldsymbol{W}_i, \boldsymbol{Z}_i | \boldsymbol{\theta}, \boldsymbol{\delta}_i, \boldsymbol{\lambda}_i, \boldsymbol{u}_i, y_{i0})$ depicts the joint probability density of the observed data for each $i \in [1, N]$, $f(\boldsymbol{\delta}_i, \boldsymbol{\lambda}_i, \boldsymbol{u}_i, y_{i0} | \boldsymbol{\theta})$ illustrates the distribution assigned in the second stage of the model for each $i \in [1, N]$, and $\pi(\boldsymbol{\theta})$ depicts the prior distribution of the unknown structural parameter vector $\boldsymbol{\theta}$.

Since, $\{\boldsymbol{\delta}_i\}_{i\in[1,N]}, \{\boldsymbol{\lambda}_i\}_{i\in[1,N]}, \boldsymbol{u}, \boldsymbol{y}_0$ are elements not observed by the researcher, the usual approach is to integrate them out of the augment posterior distribution. Therefore, the marginal posterior distribution of the structural parameters $\boldsymbol{\theta}$, can be obtained as:

$$p(\boldsymbol{\theta} | \boldsymbol{Y}, \boldsymbol{W}, \boldsymbol{Z}) \propto \int\int\int\int p(\boldsymbol{\theta}, \{\boldsymbol{\delta}_i\}_{i\in[1,N]}, \{\boldsymbol{\lambda}_i\}_{i\in[1,N]}, \boldsymbol{u}, \boldsymbol{y}_0 | \boldsymbol{Y}, \boldsymbol{W}, \boldsymbol{Z}) d\{\boldsymbol{\delta}_i\} d\{\boldsymbol{\lambda}_i\} d\boldsymbol{u} d\boldsymbol{y}_0 \tag{23}$$

The marginal posterior distribution in (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 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)[11].

5.3 The Joint Augmented Posterior

Using the above Bayesian hierarchical structure, the joint augmented posterior density of the unknown parameter vector $\boldsymbol{\Theta} = [\bar{\boldsymbol{\delta}}, \{\boldsymbol{\delta}_i\}_{i=\in[1,N]}, \bar{\boldsymbol{\lambda}}, \{\boldsymbol{\lambda}_i\}_{i\in[1,N]}, \sigma_v^2, \sigma_u^2, \sigma_e^2, \theta, \boldsymbol{\Omega}, \boldsymbol{A}, \{y_{i0}\}_{i=\in[1,N]}, \boldsymbol{u}]'$,

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

will be:

$$p(\boldsymbol{\Theta}|\boldsymbol{Y}, \boldsymbol{W}, \boldsymbol{Z}, \boldsymbol{\Psi}_0, \boldsymbol{\Phi}_0, v_0, \boldsymbol{\mu}_0, \boldsymbol{\Lambda}_0, \boldsymbol{\lambda}_0, \boldsymbol{\Gamma}_0, a_0, a_1\gamma_0, \gamma_1, h_0, h_1, b_0, b_1) \propto \tag{24}$$

$$(\sigma_v^2)^{-\frac{NT}{2}}(1-\theta^2)^{-\frac{NT}{2}}$$

$$exp\left\{-\frac{1}{2\sigma_v^2(1-\theta^2)}\sum_{i=1}^{N}(\boldsymbol{y}_i - \boldsymbol{W}_i\boldsymbol{\delta}_i - \theta\frac{\sigma_v}{\sigma_e}(\boldsymbol{q}_i - \boldsymbol{Z}_i\boldsymbol{\lambda}_i) + \boldsymbol{u}_i)'(\boldsymbol{y}_i - \boldsymbol{W}_i\boldsymbol{\delta}_i - \theta\frac{\sigma_v}{\sigma_e}(\boldsymbol{q}_i - \boldsymbol{Z}_i\boldsymbol{\lambda}_i) + \boldsymbol{u}_i)\right\}$$

$$\times (\sigma_e^2)^{-\frac{NT}{2}} exp\left\{-\frac{1}{2\sigma_e^2}\sum_{i=1}^{N}(\boldsymbol{q}_i - \boldsymbol{Z}_i\boldsymbol{\lambda}_i)'(\boldsymbol{q}_i - \boldsymbol{Z}_i\boldsymbol{\lambda}_i)\right\}$$

$$\times (\sigma_u^2)^{-\frac{NT}{2}} exp\left\{-\frac{1}{2\sigma_u^2}\sum_{i=1}^{N}\boldsymbol{u}_i'\boldsymbol{u}_i\right\}\mathbf{1}(\boldsymbol{u}_i \geq 0)$$

$$\times\, det(\boldsymbol{\Omega})^{-\frac{N}{2}} exp\left\{-\frac{1}{2}\sum_{i=1}^{N}(\boldsymbol{\delta}_i - \bar{\boldsymbol{\delta}})'\boldsymbol{\Omega}^{-1}(\boldsymbol{\delta}_i - \bar{\boldsymbol{\delta}})\right\}$$

$$\times\, det(\boldsymbol{\Lambda}_0)^{-\frac{1}{2}} exp\left\{-\frac{1}{2}(\bar{\boldsymbol{\delta}} - \boldsymbol{\mu}_0)'\boldsymbol{\Lambda}_0^{-1}(\bar{\boldsymbol{\delta}} - \boldsymbol{\mu}_0)\right\}$$

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

$$\times\, det(\boldsymbol{A})^{-\frac{N}{2}} exp\left\{-\frac{1}{2}\sum_{i=1}^{N}(\boldsymbol{\lambda}_i - \bar{\boldsymbol{\lambda}})'\boldsymbol{A}^{-1}(\boldsymbol{\lambda}_i - \bar{\boldsymbol{\lambda}})\right\}$$

$$\times\, det(\boldsymbol{\Gamma}_0)^{-\frac{1}{2}} exp\left\{-\frac{1}{2}(\bar{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_0)'\boldsymbol{\Gamma}_0^{-1}(\bar{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_0)\right\}$$

$$\times\, det(\boldsymbol{A})^{-\frac{J_2+v_0+1}{2}} exp\left\{-\frac{1}{2}tr(\boldsymbol{\Phi}_0\boldsymbol{A}^{-1})\right\}$$

$$\times\, \pi(\sigma_v^2|a_0, a_1)\pi(\sigma_u^2|\gamma_0, \gamma_1)\pi(\sigma_e^2|h_0, h_1)\pi(\theta|b_0, b_1)$$

$$\times (\sigma_0^2)^{-\frac{N}{2}} exp\left\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{N}(y_{i0} - \bar{y}_0)^2\right\}$$

The first three lines corresponds to the joint likelihood of the two endogenous variables, $\boldsymbol{y}_i$ and $\boldsymbol{q}_i$ for all $i \in [1, N]$. The fourth line determines the half-normal distribution of the technical inefficiency term $\{\boldsymbol{u}_i\}_{i\in[1,N]}$. The remaining equations depict the hierarchical structure and the prior information assigned by the researcher for the random parameters $\boldsymbol{\delta}_i$ and $\boldsymbol{\lambda}_i$, and the 'structural' parameters $\bar{\boldsymbol{\delta}}$, $\bar{\boldsymbol{\lambda}}$, σ_v^2, σ_u^2, σ_e^2, θ, $\boldsymbol{\Omega}$, $\boldsymbol{A}$, and the initial value conditions $\boldsymbol{y}_0$.

From the augmented posterior in (24), we can derive analytically all the conditional distributions of the parameters and set a Gibbs sampling algorithm to approximate the corresponding marginal distributions[12]. Most of the posterior conditional distributions are in closed form (their kernel functions belong to well know distributions) and are easy to draw form. For those in unknown form, we draw samples by incorporating the Metropolis-Hastings algorithm.

[12] In Appendix A, we present the conditional distributions of all parameters parameter, along with the Gibbs Sampler algorithm.

6 Model performance using simulated data

6.1 Data Generating Process

In this section, we present Monte Carlo simulations to evaluate the finite sample performance of the proposed MCMC algorithm. In particular, for the purpose of these simulations, we estimate a dynamic panel random coefficients production frontier model with a single exogenous input x_{it} and a single endogenous input q_{it} according to the following Data Generated Process (DGP):

$$y_{it} = a_i + \rho_i y_{it-1} + \beta_i x_{it} + \gamma_i q_{it} + v_{it} - u_{it}$$

$$q_{it} = \lambda_{0i} + \lambda_{1i} z_{it} + e_{it}$$

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

with

$$\begin{bmatrix} v_{it} \\ e_{it} \end{bmatrix} \sim \boldsymbol{\mathcal{N}} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_v^2 & \theta\sigma_v\sigma_e \\ \theta\sigma_v\sigma_e & \sigma_e^2 \end{bmatrix} \right)$$

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

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

where for simplicity we assume a diagonal variance-covariance matrix, which implies that the parameters are generated independently from each other. We perform this simulation for different settings, where we evaluate our model for different number of cross-sectional units $N = \{150, 300\}$, 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[13]. We follow the same approach to generate the exogenous instrument z_{it}. To ensure 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 setting where $\bar{\rho} = 0.5$ we choose $\sigma_\rho = 0.1$.[14]. For the β_i and γ_i coefficients, we use $\bar{\beta} = \bar{\gamma} = (1 - \bar{\rho})/2$ and $\sigma_\beta = \sigma_\gamma = 0.1$ for all the different simulations. For the coefficients λ_{0i} and λ_{1i}, we generate them from independent normal distributions, centered

[13]Such a parameterization is very common in many empirical applications that appear in the literature.

[14]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.

at $\bar{\lambda}_0 = 1$ and $\bar{\lambda}_1 = 2$, respectively, with $\sigma_0 = \sigma_1 = 0.1$. For the variance of the error term e_{it} we set $\sigma_e = 0.1$. In general, we believe that the choice of all the above parameters is realistic and can describe real empirical specifications. Specifically, we allow the long-run returns to scales to be one and therefore, the short-run returns to scales will be different depending on the level of production stagnation given by the autoregressive parameter $\bar{\rho}$. In addition, we set the correlation coefficient $\theta = 0.2$ which allows production shocks to affect the flexible input q_{it}. Moreover, for all the different simulations, in order to establish 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 $m = 10$.

Last, 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 if $\sigma_v > \sigma_u$ and $\mathcal{F} \to 0$, 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 long as σ_u is greater than σ_v, there will be enough evidence to identify the latent u_{it} term. In Table 1, we present the different simulation settings.

Table 1: Simulation Settings

Simulation Setting (I)	$\sigma_v = 0.1$	$\sigma_u = 0.2$	$\mathcal{F} = 2$	$N = \{150, 300\}$	$T = \{10, 20\}$
Simulation Setting (II)	$\sigma_v = 0.1$	$\sigma_u = 0.1$	$\mathcal{F} = 1$	$N = \{150, 300\}$	$T = \{10, 20\}$
Simulation Setting (III)	$\sigma_v = 0.05$	$\sigma_u = 0.15$	$\mathcal{F} = 3$	$N = \{150, 300\}$	$T = \{10, 20\}$
Simulation Setting (IV)	$\sigma_v = 0.04$	$\sigma_u = 0.20$	$\mathcal{F} = 5$	$N = \{150, 300\}$	$T = \{10, 20\}$

6.2 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 $\boldsymbol{\mu}_0 = \mathbf{0}_{4\times1}$ and $\boldsymbol{\Lambda}_0 = 10^3 \times \boldsymbol{I}_{4\times4}$. Similarly, we pick $\boldsymbol{\lambda}_0 = \mathbf{0}_{2\times1}$ and $\boldsymbol{\Gamma}_0 = 10^3 \times \boldsymbol{I}_{2\times2}$. This prior elicitation generates noninformative prior distributions; hence, our prior specification cannot dominate the likelihood function. For the variances of σ_v^2, σ_u^2 and σ_e^2 , we set $a_0 = a_1 = 10^{-2}$, $\gamma_0 = \gamma_1 = 10^{-2}$ and $h_0 = h_1 = 10^{-2}$. As discussed above, these hyperparameters generate the usual non-informative Jeffrey's prior distribution. For the variance-covariance matrices $\boldsymbol{\Omega}$ and $\boldsymbol{A}$, we set $v_0 = 0$, $\boldsymbol{\Psi}_0 = 10^{-6} \times \boldsymbol{I}_{4\times4}$ and $\boldsymbol{\Phi}_0 = 10^{-6} \times \boldsymbol{I}_{2\times2}$. These hyperparameters create vague prior probability densities and they are quite standard in the literature. Last, for the prior of parameter θ we set $b_0 = b_1 = 1$ which leads to the uniformative uniform prior density.

6.3 Results

In Tables 2-5 we present the posterior estimates of our simulations. In particular, for the different number of firms N and time periods T, we report the posterior average and the posterior standard deviation for the main parameter vector of interest. These are $\bar{a}, \bar{\rho}, \bar{\beta}, \bar{\gamma}$, $\bar{\lambda}_0$, $\bar{\lambda}_1$, σ_v and σ_u. In addition, for each simulation setting, we report the Pearson correlation between the true (the generated) and the estimated inefficiencies. The Bayesian MCMC is based on 5,000 iterations from which the first 1,500 posterior draws are discarded from our analysis to eliminate potential effects of the initial values.

[Table 2 here]

[Table 3 here]

[Table 4 here]

[Table 5 here]

Overall, we see that for all different simulation settings, 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, including the input elasticity $\bar{\gamma}$ of the endogenous input q_{it}. In addition, we see that in all settings, we are able to obtain the correct posterior average for σ_u, which highlights that we are able to estimate very well the average of the latent technical inefficiency.

Furthermore, 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 Simulation Setting (I) presented in Table 2, where the $\mathcal{F} = 2$ we see that the correlation coefficient is around 0.69 and increases to 0.74, as T increases from 10 to 20. On the other hand, as $\mathcal{F}$ increases to 3 and 5, as presented in Simulation Setting (III) and (IV) in Tables 4 and 5, respectively, we observe that the correlation coefficient increases from around 0.76 and 0.84, for $T = 10$ and $T = 20$, to 0.84 and 0.90, for $T = 10$ and $T = 20$, respectively. On the contrary, as we can see from Table 3 where the signal-to-noise ratio $\mathcal{F} = 1$, the correlation between the true and estimated inefficiencies drop to a range between 0.40 and 0.50, 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.

7 Empirical Study

7.1 Data

In this section, we demonstrate the new model using a dataset of publicly listed UK companies operating in the manufacturing and construction sector covering the period 2004-2022. The data are collected from the Financial Analysis Made Easy (FAME) database provided by Bureau Van Bijk Moody's Analytics[15]. The database covers over 11 million companies in the UK and Ireland and includes detailed financial statements, including companies' balance sheets and income statements. The database has been used in many other productivity studies such as in Harris and Li (2008), Guariglia and Mateut (2010), Draca et al. (2011), Gong and Sickles (2020), among many others.

For the model illustration, we focus only on publicly traded companies operating in the manufacturing and construction sectors. We select these two sectors because the production process is highly connected with new equipment installation and machinery upgrades, workforce training and retraining, which are quasi-fixed. Moreover, we use data only for public companies, since our choice is driven by the fact that we are interested in illustrating the new specification rather than drawing inferences for the entire population of firms in the two sectors. In addition, we want to minimize any problems arising from extensive heterogeneity, missing data and other data-related issues that naturally manifest when dealing with firm-level data. After dropping all raw data that exhibit negative values and keeping companies with available data for at least 7 consecutive years, we were able to identify 183 large companies that are traded on London Stock Exchange[16]. The final dataset consists of an unbalanced panel with 3,115 data observations and the yearly observations for each firm range from 7 to 19.

For the empirical application, we assume a production technology where firms use capital, labour and intermediate inputs in order to produce a single output. We allow capital and labour to be quasi-fixed, while intermediate expenditure is variable. This implies that intermediate expenditure could be correlated with current production shocks.

We construct the variables following the method as in Gong and Sickles (2020). Specifically, we have: for output we use the Total Turnover, for capital we use the "Capital Employed" defined as the Total Assets less the Current Liabilities. Labour is defined as the total number of employees. Since the database does not provide direct estimates of the intermediate cost, we construct the intermediate input as the Total Turnover less the Total Value Added. We estimate Total Value Added as the sum of Operating Profits, Capital and Labour Expenditures.

[15] https://www.bvdinfo.com/en-gb/our-products/data/national/fame

[16] In Appendix B, Table 9, we present the two-digit UK SIC (2007) codes of the firms in our sample.

Last, Capital Expenditure is defined as the sum of depreciation, amortization and impairments, and Labour Expenditure is the remuneration[17]. In addition, we use the Output Producer Price Index to deflate the Turnover and the Capital Employed and the Input Producer Price Index to deflate the Intermediate expenditure. Last, we use the UK Economic Policy Uncertainty Index[18] (EPU Index) to capture non-linear productivity shocks.

In the next table, we present some descriptive statistics.

[Table 6 here]

Last, given the potential endogeneity problem related to intermediate expenditure, we instrument this variable using global prices of different commodities. The rationale is that commodity prices are perpendicular to production shocks and are highly correlated with our endogenous input. This is based on the intuition that individual firms are unable to influence global commodity prices and thus behave as price takers. Moreover, it is important to note that our measure of intermediate expenditure primary comprises material expenditure and transportation costs, both of which are highly correlated with the global prices used as instruments. In addition, we use the Global Supply Chain Pressure Index (GSCPI) developed by the Federal Reserve Bank of New York to account for global supply chain distortions. Here, we need to highlight that the final instruments employed are not the raw prices themselves but rather indices derived from these prices, a procedure that we explain in detail in the next subsection.

In the next table, we present some descriptive statistics of all the instrumental variables we use. The table shows the prices in USD. Before we proceed with our analysis, we convert them to GBP in order to be consistent with the rest of the variables, and we deflate them using the UK CPI to reflect real monetary values.

[Table 7 here]

7.2 The empirical model

7.2.1 DRC-SFM

For the empirical model we opt for a Cobb-Douglas function for the long-run specification. Hence, the corresponding dynamic specification (DRC-SFM) will be:

$$y_{it} = a_i + \rho_i \, y_{it-1} + \beta_{ki} \, k_{it} + \beta_{li} \, l_{it} + \beta_{mi} \, m_{it} + \beta_{ti} t + \beta_{UIi} epu_t + v_{it} - u_{it} \tag{25}$$

[17] Remuneration includes Wages and Salaries, Social Security costs, Pension costs and Other Staff costs. For this study, we opt not to include the director's remuneration in the labour expenditure, as it will highly inflate the cost of labour.

[18] The index is based on work presented at Baker et al. (2016). The dataset is available at `https://www.policyuncertainty.com/index.html`.

where y_{it}, k_{it}, l_{it} and m_{it} are the output, capital, labour, as well the intermediate inputs used in each firm i at time t. In the analysis, all variables are used in natural logarithm form. In addition, to account for the firm-specific technological progress and the common macroeconomic and productivity effects, we include a time trend t and the natural logarithm of EPU index epu_t, to capture nonlinear productivity shocks. From the above empirical specification, the structure of firm-specific parameters will be:

$$\boldsymbol{\delta}_i \equiv [a_i, \rho_i, \beta_{ki}, \beta_{li}, \beta_{mi}, \beta_{ti}, \beta_{UIi}]' \sim \mathcal{N}(\bar{\boldsymbol{\delta}}, \boldsymbol{\Omega})$$

where $\bar{\boldsymbol{\delta}} = [\bar{a}, \bar{\rho}, \bar{\beta}_k, \bar{\beta}_l, \bar{\beta}_m, \bar{\beta}_t, \bar{\beta}_{UI}]'$ and $\boldsymbol{\Omega}$ is the corresponding variance-covariance matrix. Last, we allow production shocks and the non-negative inefficiency effects to exhibit firm-specific heteroscedasticity, viz. $v_{it} \sim \mathcal{N}(0, \sigma^2_{vi})$ and $u_{it} \sim \mathcal{N}^+(0, \sigma^2_{ui})$ for all $i = 1, 2, \ldots, N$, as our purpose is to allow the model to be as flexible as possible.

7.2.2 DRC-E-SFM

Extending the DRC-SFM, we estimate the proposed DRC-E-SFM, by allowing the intermediate input m_{it} to be correlated with production shocks, viz. $Cov(m_{it}; v_{it}) \neq 0$. We couple the main equation, with the following:

$$m_{it} = \boldsymbol{z}'_{it}\boldsymbol{\lambda}_i + e_{it} \tag{26}$$

where $\boldsymbol{z}_{it}$ is a vector of valid instruments. In addition, we allow the structure between v_{it} and e_{it} to have the following form:

$$\begin{bmatrix} v_{it} \\ e_{it} \end{bmatrix} \sim \boldsymbol{\mathcal{N}}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_v & \theta\sigma_v\sigma_e \\ \theta\sigma_v\sigma_e & \sigma^2_e \end{bmatrix}\right)$$

To derive appropriate instruments we use all variables in Table 7 in real GBP values. As many of them are nonstationary in nature, we compute the year-on-year percentage changes. First, as Figure 1 indicates, many of the variables are highly correlated with each otherx. For this reason, we proceed with Principal Components Analysis in order to (i) reduce dimensionality of this dataset and (ii) solve any potential multicollinearity issues. Figure 2 illustrates the cumulative variance explained by all the principal components. We can see that the first three components, viz. PC1, PC2 and PC3, explain more than 80% of the total variation. Thus, we use the first three components as appropriate instrumental variables. In addition, we include a time trend to capture the general trend of each intermediate expenditure. Thus, PC1, PC2 and PC3 are used to capture deviations from the long-run trend.

[Figure 1 here]

[Figure 2 here]

Hence, the final form of the equation (26) will be:

$$m_{it} = \lambda_{0i} + \lambda_{1i}PC1_{it} + \lambda_{2i}PC2_{it} + \lambda_{3i}PC3_{it} + \lambda_{4i}t + e_{it} \tag{27}$$

where

$$\boldsymbol{\lambda}_i \equiv [\lambda_{0i}, \lambda_{1i}, \lambda_{2i}, \lambda_{3i}, \lambda_{4i}]' \sim \mathcal{N}(\bar{\boldsymbol{\lambda}}, \boldsymbol{A})$$

with $\bar{\boldsymbol{\lambda}} = [\bar{\lambda}_0, \bar{\lambda}_1, \bar{\lambda}_2, \bar{\lambda}_3, \bar{\lambda}_4]'$ and $\boldsymbol{A}$ to be the corresponding variance covariance matrix[19].

7.3 Prior Specifications

Before we proceed with posterior model analysis, we present the prior elicitation for the DRC-SFM and DRC-E-SFM. For the variance of the error term σ_{vi}^2 we set $a_0 = a_1 = 10^{-2}$ for all cross-sectional units. This leads to non-informative Inverse-Gamma priors. We use the same priors for the inefficiency variances σ_{ui}^2 and we set $\gamma_0 = \gamma_1 = 10^{-2}$.

For the random coefficients $\boldsymbol{\delta}_i$, we set $\boldsymbol{\mu}_0 = 0 \times \mathbf{1}_{7\times1}$ and $\Lambda_0 = 100 \times \boldsymbol{I}_{7\times7}$. This generates non-informative prior densities for the parameters of the main equation (25). We follow the same approach for the parameter vector $\boldsymbol{\lambda}_i$ and we set $\boldsymbol{\lambda}_0 = 0 \times \mathbf{1}_{5\times1}$ and $\boldsymbol{\Gamma}_0 = 100 \times \boldsymbol{I}_{5\times5}$.

For the variance-covariance matrix $\boldsymbol{\Omega}$ we choose $v_0 = 0$ and $\boldsymbol{\Psi}_0 = 10^{-6} \times \boldsymbol{I}_{7\times7}$. Similarly, for $\boldsymbol{A}$ we set $v_0 = 0$ and $\boldsymbol{\Phi}_0 = 10^{-6} \times \boldsymbol{I}_{5\times5}$. These generate the non-informative Jeffrey's prior distribution for all variance covariance matrices.

For the variance σ_e^2 we set $h_0 = h_1 = 10^{-2}$. For parameter θ we set $b_0 = b_1 = 1$ which produced a vague uniform distribution on the support $[0, 1]$. Last, for the latent initial values, we assume $y_{i0} \sim \mathcal{N}(y_{i1}, \sigma_0^2)$ and we set $\sigma_0 = 0.2$.

7.4 Empirical Results

The next table presents the posterior averages and standard deviations. For both models, all Gibbs sampling computations are performed using 5,000 iterations from which the first 30% of the draws are discarded from our analysis to mitigate any initial value effect. Therefore, the empirical inference is based on the remaining 3,500 posterior sample draws.

[Table 8 here]

First, from Table 8, we compare the results between the DRC-SFM and the DRC-E-SFM that allows intermediate input to be flexible, and thus correlated with the production shocks v_{it}.

[19] In Appendix C, we present a framework for testing the relevance of the proposed instruments.

Overall, both models generate very similar posterior averages, indicating that in our empirical study, the omitted variable bias is negligible. This can also be verified from the fact that for the parameter θ (correlation between v_{it} and e_{it}), the posterior average is 0.070, with posterior std of 0.034.

Turning our attention to the estimated technologies, it is interesting to highlight that the estimated Returns to Scales (RTS), are 0.781 and 0.824, respectively. This demonstrates that in the short run, firms face upward sloping average cost curves reflecting the fact that production inputs do not contribute instantaneously at their full capacities.

Regarding the average technical efficiencies, we see that both DRC-SFM and DRC-E-SFM generate very similar estimated technical efficiencies. Specifically, the average efficiency estimates are 90% and 89%, respectively, indicating that firms, although they operate near their frontier, could potentially increase their output through efficiency improvement practices[20]. Figures 3 and 4 illustrate technical efficiency estimates for both models. It is evident that overall the technical estimates are very similar between the two models. However, it is important to note that, when investigating the efficiency differences across the entire distribution, the model that accounts for input endogeneity generates slightly lower efficiency scores. This is due to the fact that the DRC-E-SFM generates lower autoregressive parameters, and hence, the output fall is attributed to technical inefficiency rather than the inherent stagnant behaviour of the production process. This is an important result, as it highlights a fundamental distinction: the interpretation of output shortfalls depends on whether they are attributed to intertemporal adjustment frictions or to inefficiencies in production. In models with higher production persistence (larger $\bar{\rho}$), output deviations are more likely to be interpreted as a result of sluggish adjustment. In contrast, when endogeneity is accounted for and $\bar{\rho}$ decreases, the same deviations are more likely to be seen as technical inefficiencies.

[Figure 3 here]

[Figure 4 here]

Regarding the evolution of the efficiency scores over time, Figure 5 illustrates the trend in average efficiency for both the DRC-SFM and DRC-E-SFM. Overall, both specifications indicate that the technical efficiency followed a three-stage cycle. First, during the period between 2004 and 2012, the average efficiency decreased over time. However, in the following years, up the 2016, firms managed to recover and considerably improve their efficiency scores. Lastly, in the

[20]In Appendix D, we conduct robustness analysis of the efficiency estimates. Specifically, we re-estimate the model using different inefficiency prior distributions.

post-2016 period, which was highly turbulent for U.K. and marked by significant events such as the Brexit referendum, the Brexit itself, and, of course, the COVID-19 pandemic, average efficiency scores decreased by approximately 5 percentage points. In addition, both models provide evidence that during the 2004-2022 the technical change has been negative.

[Figure 5 here]

Last, we see that both models generate significant estimates for the autoregressive parameter. Overall, this points to the fact that the production processes are stagnant, and therefore, a dynamic specification seems more plausible than a simple static specification. Regarding the estimated $\bar{\rho}$, the posterior averages from the DRC-SFM and DRC-E-SFM are 0.214 and 0.168, respectively. However, it is important to highlight the fact that results across the different firms are quite heterogeneous. In Figure 6, we illustrate the empirical densities of each estimated autoregressive parameter ρ_i. From the graph, we observed that despite some firms being able to adjust their production relatively fast (autoregressive parameters less than 0.1), the majority of firms exhibit mild to subsequent stagnant production behaviours. Specifically, the firm-specific autoregressive parameters can be as large as 0.70, indicating that most firms faced substantial challenges in adjusting their production processes. This can be attributed to the fact that companies encountered two major macroeconomic shocks during the period of 2004-2022, namely the financial crisis of 2008-2009 and the COVID-19 crisis of 2020-2022. These two major events created considerable macroeconomic uncertainties which affected the firms decision regarding their production expansionary or contractionary policies. Moreover, the period of our analysis is characterized as a period of green transition, where firms tried to adopt new greener technologies in order to comply with a set of mandatory green policies and regulations, such as the EU Emissions Trading System during the initial phase, and the UK ETS at the later stage. Consequently, the adjustment process of transitioning could be quite costly and sluggish. Regarding the time period required for the production process to catch up with the desired level, we compute the "half-life" which denotes the time period that is needed for the production to move halfway towards the targeted level[21]. We find that the average half-life is 5 months and can be up to 6 quarters.

[Figure 6 here]

[21] Given an estimate for ρ_i, half-life can be computed as $ln(0.5)/ln(\rho_i)$.

8 Conclusions

This paper introduces a novel Bayesian dynamic panel stochastic frontier model that accounts for partial output adjustment, heterogeneous technologies, and time-varying technical inefficiency while addressing input endogeneity through a joint modeling framework. By incorporating a partial adjustment mechanism and allowing for firm-specific dynamics, the model captures realistic frictions in production processes that are often ignored in static frontier models.

Estimation is conducted within a Bayesian hierarchical structure using Markov Chain Monte Carlo (MCMC) techniques, enabling efficient inference even in panels of small and moderate size. Through extensive simulation exercises, we demonstrate that the model performs well in recovering structural parameters and inefficiency scores under various settings, particularly when the signal-to-noise ratio is sufficient.

Applying the model to a panel of UK publicly listed manufacturing and construction firms over 2004–2022, we uncover strong evidence of production inertia and adjustment costs, with many firms exhibiting sluggish responses to production shocks. The estimated average technical efficiency is approximately 89%, suggesting that while firms operate relatively close to their production frontiers, substantial gains remain possible through improved operational practices. Furthermore, the analysis highlights the importance of accounting for macroeconomic shocks and regulatory transitions, such as the global financial crisis, Brexit, and the green transition, which have had significant impacts on efficiency dynamics.

The findings carry important policy implications. Specifically, during periods of economic turbulence or structural change, policymakers should recognize the existence of adjustment frictions and inefficiencies, and avoid implementing overly rigid policies that could exacerbate cost pressures on firms. Flexible, supportive regulatory environments are more conducive to enabling firms to transition towards more productive and sustainable practices.

Acknowledgements

Sadly, Mike Tsionas passed away in 2024. The surviving authors dedicate this paper to his memory. We are grateful to Alessio Sancetta, Geraint Johnes, Subal Kumbhakar, Byunghoon Kang, Ruijun Bu, participants at NWSSDTP conference at Manchester University and at Lancaster University Economics seminars for their valuable comments. Any errors are our own.

References

Ackerberg, D. A., Caves, K., and Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, 83(6):2411–2451.

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.

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.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4):1593–1636.

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.

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

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*.

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.

Draca, M., Machin, S., and Van Reenen, J. (2011). Minimum wages and firm profitability. *American Economic Journal: Applied Economics*, 3(1):129–151.

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

Emvalomatis, G. (2020). Semi-parametric analysis of efficiency and productivity using gaussian processes. *The Econometrics Journal*, 23(1):48–67.

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 Greene, W. (2016). Persistent and transient productive inefficiency: a maximum simulated likelihood approach. *Journal of Productivity Analysis*, 45:187–196.

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.

Gandhi, A., Navarro, S., and Rivers, D. A. (2020). On the identification of gross output production functions. *Journal of Political Economy*, 128(8):2973–3016.

Gong, B. and Sickles, R. C. (2020). Non-structural and structural models in productivity analysis: study of the british isles during the 2007–2009 financial crisis. *Journal of Productivity Analysis*, 53:243–263.

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.

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

Guariglia, A. and Mateut, S. (2010). Inventory investment, global engagement, and financial constraints in the UK: evidence from micro data. *Journal of Macroeconomics*, 32(1):239–250.

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

Harris, R. and Li, Q. C. (2008). Evaluating the contribution of exporting to UK productivity growth: some microeconomic evidence. *World Economy*, 31(2):212–235.

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*.

Kim, Y. and Kim, C.-J. (2011). Dealing with endogeneity in a time-varying parameter model: joint estimation and two-step estimation procedures. *The Econometrics Journal*, 14(3):487–497.

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

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.

Levinsohn, J. and Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The review of economic studies*, 70(2):317–341.

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., 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.

Olley, G. S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica: Journal of the Econometric Society*, pages 1263–1297.

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

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

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.

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, E. G. and Kumbhakar, S. C. (2014). Firm heterogeneity, persistent and transient technical inefficiency: A generalized true random-effects model. *Journal of Applied Econometrics*, 29(1):110–132.

Tsionas, M., Parmeter, C. F., and Zelenyuk, V. (2023). Bayesian artificial neural networks for frontier efficiency analysis. *Journal of Econometrics*, 236(2):105491.

Tsionas, M. G. (2017). “When, where, and how” of efficiency estimation: improved procedures for stochastic frontier modeling. *Journal of the American Statistical Association*, 112(519):948–965.

Tsionas, M. G. and Kumbhakar, S. C. (2023). Productivity and performance: A GMM approach. *Oxford Bulletin of Economics and Statistics*, 85(2):331–344.

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.

Table 2: Posterior Estimates for the Simulation Setting (I)

Panel A

		$N = 150$				$N = 300$			
		$T = 10$		$T = 20$		$T = 10$		$T = 20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	0.975	0.012	0.196	0.011	0.995	0.009	0.984	0.008
$\bar{\rho}$	0.20	0.194	0.004	0.196	0.005	0.202	0.005	0.201	0.004
$\bar{\beta}$	0.40	0.390	0.009	0.407	0.009	0.408	0.007	0.408	0.006
$\bar{\gamma}$	0.40	0.399	0.007	0.385	0.009	0.405	0.007	0.409	0.006
$\bar{\lambda}_0$	1.00	1.001	0.009	0.995	0.009	1.008	0.006	1.001	0.006
$\bar{\lambda}_1$	2.00	2.000	0.009	1.996	0.008	1.996	0.007	2.008	0.006
σ_v	0.10	0.121	0.008	0.103	0.003	0.100	0.002	0.104	0.001
σ_u	0.20	0.162	0.016	0.200	0.007	0.185	0.007	0.194	0.004
$cor(u_{it}; \hat{u}_{it})$		0.682		0.736		0.646		0.732	

Panel B

		$N = 150$				$N = 300$			
		$T = 10$		$T = 20$		$T = 10$		$T = 20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	1.021	0.010	0.986	0.007	0.994	0.004	0.981	0.010
$\bar{\rho}$	0.50	0.485	0.003	0.498	0.004	0.501	0.003	0.503	0.005
$\bar{\beta}$	0.25	0.239	0.009	0.257	0.009	0.257	0.007	0.258	0.006
$\bar{\gamma}$	0.25	0.249	0.007	0.235	0.009	0.254	0.006	0.259	0.006
$\bar{\lambda}_0$	1.00	1.002	0.009	0.995	0.009	1.008	0.006	1.001	0.006
$\bar{\lambda}_1$	2.00	2.000	0.009	1.996	0.008	1.995	0.007	2.007	0.006
σ_v	0.10	0.101	0.003	0.107	0.002	0.107	0.002	0.102	0.002
σ_u	0.20	0.191	0.009	0.194	0.006	0.180	0.006	0.198	0.005
$cor(u_{it}; \hat{u}_{it})$		0.690		0.739		0.650		0.736	

Panel C

		$N = 150$				$N = 300$			
		$T = 10$		$T = 20$		$T = 10$		$T = 20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	0.780	0.004	1.004	0.007	0.765	0.011	0.926	0.003
$\bar{\rho}$	0.80	0.847	0.004	0.789	0.004	0.856	0.004	0.814	0.003
$\bar{\beta}$	0.10	0.091	0.009	0.107	0.009	0.110	0.007	0.108	0.006
$\bar{\gamma}$	0.10	0.101	0.008	0.085	0.009	0.105	0.006	0.109	0.006
$\bar{\lambda}_0$	1.00	1.001	0.009	0.995	0.009	1.008	0.006	1.001	0.006
$\bar{\lambda}_1$	2.00	2.000	0.009	1.996	0.008	1.995	0.007	2.007	0.006
σ_v	0.10	0.103	0.007	0.114	0.004	0.117	0.004	0.101	0.001
σ_u	0.20	0.193	0.014	0.182	0.009	0.101	0.008	0.200	0.004
$cor(u_{it}; \hat{u}_{it})$		0.688		0.738		0.642		0.741	

Notes: The table reports the posterior averages and standard deviations of the main parameters of interest $\{\bar{a}, \bar{\rho}, \bar{\beta}, \bar{\gamma}, \bar{\lambda}_0, \bar{\lambda}_1, \sigma_v, \sigma_u\}$. The MCMC algorithm is based on 5,000 iterations from which the first 1,500 samples are discarded. $cor(u_{it}; \hat{u}_{it})$ denotes the Pearson's correlation between the estimated $\hat{u}_{it}$ and generated u_{it} inefficiencies.

Table 3: Posterior Estimates for the Simulation Setting (II)

Panel A

		$N = 150$				$N = 300$			
		$T = 10$		$T = 20$		$T = 10$		$T = 20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	0.993	0.005	0.989	0.011	0.977	0.011	0.978	0.010
$\bar{\rho}$	0.20	0.194	0.007	0.189	0.008	0.203	0.007	0.201	0.006
$\bar{\beta}$	0.40	0.389	0.008	0.408	0.009	0.406	0.006	0.408	0.006
$\bar{\gamma}$	0.40	0.399	0.007	0.384	0.009	0.404	0.006	0.409	0.006
$\bar{\lambda}_0$	1.00	1.001	0.009	0.996	0.009	1.008	0.006	1.001	0.006
$\bar{\lambda}_1$	2.00	2.000	0.009	1.995	0.008	1.996	0.007	2.007	0.006
σ_v	0.10	0.110	0.003	0.107	0.002	0.109	0.002	0.105	0.003
σ_u	0.10	0.087	0.010	0.090	0.008	0.063	0.01	0.085	0.009
$cor(u_{it}; \hat{u}_{it})$		0.398		0.491		0.392		0.478	

Panel B

		$N = 150$				$N = 300$			
		$T = 10$		$T = 20$		$T = 10$		$T = 20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	0.995	0.009	0.990	0.011	1.001	0.004	0.986	0.009
$\bar{\rho}$	0.50	0.500	0.005	0.496	0.004	0.502	0.003	0.502	0.004
$\bar{\beta}$	0.25	0.240	0.008	0.258	0.009	0.256	0.007	0.258	0.006
$\bar{\gamma}$	0.25	0.249	0.007	0.235	0.008	0.254	0.006	0.259	0.006
$\bar{\lambda}_0$	1.00	1.001	0.009	0.995	0.009	1.008	0.006	1.001	0.006
$\bar{\lambda}_1$	2.00	2.000	0.009	1.996	0.008	1.995	0.007	2.008	0.006
σ_v	0.10	0.099	0.004	0.103	0.002	0.103	0.002	0.100	0.001
σ_u	0.10	0.096	0.013	0.098	0.007	0.091	0.007	0.100	0.005
$cor(u_{it}; \hat{u}_{it})$		0.426		0.498		0.402		0.482	

Panel C

		$N = 150$				$N = 300$			
		$T = 10$		$T = 20$		$T = 10$		$T = 20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	0.835	0.004	0.976	0.007	0.900	0.004	0.907	0.015
$\bar{\rho}$	0.80	0.828	0.004	0.797	0.005	0.820	0.003	0.812	0.004
$\bar{\beta}$	0.10	0.092	0.009	0.108	0.009	0.107	0.007	0.108	0.006
$\bar{\gamma}$	0.10	0.100	0.008	0.085	0.009	0.104	0.006	0.109	0.006
$\bar{\lambda}_0$	1.00	1.001	0.009	0.996	0.009	1.008	0.006	1.001	0.006
$\bar{\lambda}_1$	2.00	2.000	0.009	1.996	0.008	1.996	0.007	2.007	0.006
σ_v	0.10	0.108	0.003	0.099	0.003	0.104	0.002	0.108	0.001
σ_u	0.10	0.071	0.012	0.107	0.008	0.082	0.008	0.075	0.006
$cor(u_{it}; \hat{u}_{it})$		0.427		0.502		0.405		0.489	

Notes: The table reports the posterior averages and standard deviations of the main parameters of interest $\{\bar{a}, \bar{\rho}, \bar{\beta}, \bar{\gamma}, \bar{\lambda}_0, \bar{\lambda}_1, \sigma_v, \sigma_u\}$. The MCMC algorithm is based on 5,000 iterations from which the first 1,500 samples are discarded. $cor(u_{it}; \hat{u}_{it})$ denotes the Pearson's correlation between the estimated $\hat{u}_{it}$ and generated u_{it} inefficiencies.

Table 4: Posterior Estimates for the Simulation Setting (III)

Panel A

		$N = 150$				$N = 300$			
		$T = 10$		$T = 20$		$T = 10$		$T = 20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	1.011	0.011	0.994	0.010	0.995	0.007	0.990	0.006
$\bar{\rho}$	0.20	0.192	0.007	0.188	0.008	0.203	0.007	0.200	0.003
$\bar{\beta}$	0.40	0.386	0.008	0.407	0.009	0.407	0.006	0.407	0.006
$\bar{\gamma}$	0.40	0.398	0.007	0.385	0.009	0.404	0.006	0.409	0.006
$\bar{\lambda}_0$	1.00	1.001	0.009	0.995	0.009	1.008	0.006	1.001	0.006
$\bar{\lambda}_1$	2.00	2.000	0.009	1.996	0.008	1.995	0.007	2.007	0.006
σ_v	0.05	0.050	0.002	0.055	0.003	0.053	0.001	0.052	0.001
σ_u	0.15	0.150	0.005	0.145	0.005	0.137	0.004	0.149	0.002
$cor(u_{it}; \hat{u}_{it})$		0.758		0.825		0.719		0.823	

Panel B

		$N = 150$				$N = 300$			
		$T = 10$		$T = 20$		$T = 10$		$T = 20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	1.015	0.008	0.996	0.012	0.992	0.010	0.989	0.008
$\bar{\rho}$	0.50	0.494	0.006	0.488	0.009	0.506	0.007	0.501	0.006
$\bar{\beta}$	0.25	0.236	0.008	0.258	0.009	0.257	0.006	0.257	0.006
$\bar{\gamma}$	0.25	0.248	0.007	0.234	0.008	0.254	0.006	0.259	0.006
$\bar{\lambda}_0$	1.00	1.002	0.009	0.995	0.009	1.008	1.995	1.001	0.006
$\bar{\lambda}_1$	2.00	2.000	0.009	1.996	0.008	0.006	0.007	2.007	0.006
σ_v	0.05	0.052	0.002	0.053	0.001	0.049	0.002	0.050	0.001
σ_u	0.15	0.149	0.005	0.148	0.004	0.140	0.003	0.151	0.002
$cor(u_{it}; \hat{u}_{it})$		0.764		0.831		0.730		0.826	

Panel C

		$N = 150$				$N = 300$			
		$T = 10$		$T = 20$		$T = 10$		$T = 20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	0.962	0.012	0.992	0.009	0.904	0.004	0.957	0.015
$\bar{\rho}$	0.80	0.804	0.005	0.794	0.004	0.822	0.003	0.807	0.004
$\bar{\beta}$	0.10	0.089	0.008	0.107	0.009	0.108	0.006	0.107	0.006
$\bar{\gamma}$	0.10	0.099	0.007	0.084	0.009	0.104	0.006	0.109	0.006
$\bar{\lambda}_0$	1.00	1.001	0.009	0.995	0.009	1.008	0.006	1.001	0.006
$\bar{\lambda}_1$	2.00	2.000	0.009	1.996	0.008	1.996	0.007	2.008	0.006
σ_v	0.05	0.055	0.002	0.052	0.001	0.051	0.001	0.051	0.001
σ_u	0.15	0.145	0.005	0.148	0.003	0.142	0.003	0.151	0.002
$cor(u_{it}; \hat{u}_{it})$		0.781		0.833		0.738		0.835	

Notes: The table reports the posterior averages and standard deviations of the main parameters of interest $\{\bar{a}, \bar{\rho}, \bar{\beta}, \bar{\gamma}, \bar{\lambda}_0, \bar{\lambda}_1, \sigma_v, \sigma_u\}$. The MCMC algorithm is based on 5,000 iterations from which the first 1,500 samples are discarded. $cor(u_{it}; \hat{u}_{it})$ denotes the Pearson's correlation between the estimated $\hat{u}_{it}$ and generated u_{it} inefficiencies.

Table 5: Posterior Estimates for the Simulation Setting (IV)

Panel A

		$N = 150$				$N = 300$			
		$T = 10$		$T = 20$		$T = 10$		$T = 20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	1.012	0.012	1.001	0.009	0.994	0.008	0.992	0.007
$\bar{\rho}$	0.20	0.192	0.008	0.187	0.008	0.204	0.007	0.200	0.006
$\bar{\beta}$	0.40	0.386	0.009	0.408	0.009	0.408	0.007	0.406	0.006
$\bar{\gamma}$	0.40	0.399	0.007	0.385	0.008	0.403	0.006	0.410	0.006
$\bar{\lambda}_0$	1.00	1.001	0.009	0.995	0.009	1.008	0.006	1.001	0.006
$\bar{\lambda}_1$	2.00	2.000	0.009	1.996	0.008	1.995	0.007	2.007	0.006
σ_v	0.04	0.037	0.001	0.041	0.001	0.043	0.001	0.042	0.001
σ_u	0.20	0.202	0.005	0.200	0.003	0.186	0.004	0.201	0.002
$cor(u_{it}; \hat{u}_{it})$		0.836		0.897		0.802		0.900	

Panel B

		$N = 150$				$N = 300$			
		$T = 10$		$T = 20$		$T = 10$		$T = 20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	1.039	0.011	1.014	0.008	0.987	0.011	0.991	0.008
$\bar{\rho}$	0.50	0.481	0.005	0.488	0.004	0.505	0.005	0.501	0.004
$\bar{\beta}$	0.25	0.236	0.009	0.258	0.009	0.258	0.006	0.256	0.006
$\bar{\gamma}$	0.25	0.248	0.007	0.234	0.009	0.254	0.006	0.260	0.006
$\bar{\lambda}_0$	1.00	1.001	0.009	0.995	0.009	1.008	0.006	1.001	0.006
$\bar{\lambda}_1$	2.00	2.000	0.009	1.996	0.008	1.995	0.007	2.007	0.006
σ_v	0.04	0.041	0.003	0.044	0.002	0.045	0.001	0.040	0.001
σ_u	0.20	0.199	0.005	0.201	0.004	0.185	0.004	0.202	0.002
$cor(u_{it}; \hat{u}_{it})$		0.853		0.899		0.816		0.897	

Panel C

		$N = 150$				$N = 300$			
		$T = 10$		$T = 20$		$T = 10$		$T = 20$	
Parameters	True Value	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std	Post. Mean	Post. Std
$\bar{a}$	1.00	0.951	0.011	0.975	0.004	0.892	0.005	0.961	0.017
$\bar{\rho}$	0.80	0.808	0.005	0.797	0.004	0.827	0.003	0.806	0.005
$\bar{\beta}$	0.10	0.088	0.009	0.107	0.009	0.109	0.007	0.106	0.006
$\bar{\gamma}$	0.10	0.100	0.007	0.084	0.009	0.104	0.006	0.110	0.006
$\bar{\lambda}_0$	1.00	1.001	0.009	0.996	0.009	1.008	0.006	1.001	0.006
$\bar{\lambda}_1$	2.00	2.000	0.009	1.995	0.008	1.995	0.007	2.008	0.006
σ_v	0.04	0.043	0.001	0.044	0.002	0.043	0.001	0.041	0.001
σ_u	0.20	0.196	0.005	0.196	0.004	0.189	0.004	0.201	0.002
$cor(u_{it}; \hat{u}_{it})$		0.857		0.907		0.818		0.904	

Notes: The table reports the posterior averages and standard deviations of the main parameters of interest $\{\bar{a}, \bar{\rho}, \bar{\beta}, \bar{\gamma}, \bar{\lambda}_0, \bar{\lambda}_1, \sigma_v, \sigma_u\}$. The MCMC algorithm is based on 5,000 iterations from which the first 1,500 samples are discarded. $cor(u_{it}; \hat{u}_{it})$ denotes the Pearson's correlation between the estimated $\hat{u}_{it}$ and generated u_{it} inefficiencies.

Table 6: Descriptive Statistics

Variable	Description	Mean	Median	Min	Max
Output (Y)	Total Turnover (in thousands of £)	1,301,689	113,208	10	36,724,000
Capital (K)	Capital Employed (in thousands of £)	1,544,015	79,793	79	135,693,000
Labour (L)	Number of Employees	5,555	748	2	113,667
Intermediate Input (M)	Turnover Less VA (in thousands of £)	771,894	63,334	114	28,288,000
PPI O	Producers Price Index Output	0.9622	1.0000	0.6603	1.4827
PPI I	Producers Price Index Input	0.9479	1.0000	0.7311	1.3204
CPI	Consumer Price Index	0.844	0.814	0.551	1.205
UK EPU Index	UK Economic Policy Uncertainty Index	130.63	132.58	47.92	266.43

Notes: The firm level data have been obtained from the FAME Database. The price indices have been obtained from the Office for National Statistics (ONS). The UK EPU Index has been obtained from `https://www.policyuncertainty.com/index.html`.

Table 7: Energy, Material and Supply Chain Indices

Commodities & Supply Indices	Description	Mean	Median	Max	Min
Energy Related Commodities					
Crude Oil	USD per Barrel (Brent-Europe)	73.61	70.86	111.63	38.26
Global Price of Natural Gas (EU)	USD per BTU	9.83	8.21	37.52	3.18
Global Price of Coal	USD per metric ton	103.08	82.19	384.17	50.82
Material Related Commodities					
Global Price of Copper	USD per metric ton	6550.28	6731.35	9317.41	2863.47
Global Price of Metal	Index in USD	143.01	140.01	209.36	67.23
Global Price of Aluminum	USD per metric ton	2074.34	1967.65	2706.99	1604.18
Global Price of Rubber	US cents per pound	101.00	88.75	218.51	58.12
Global Price of Tin	USD per metric ton	18771.31	18661.16	32387.29	7385.25
Global Price of Iron	USD per metric ton	87.81	79.99	167.79	16.39
Supply Related Index					
GSCPIC Index	Standard deviations from Mean	0.215	-0.155	3.032	-0.670

Notes: Energy and material related data has been obtained from the FRED Database. GSCPIC Index has been obtained from the Federal Reserve Bank of New York.

Table 8: Posterior estimates of the parameters

Description	Parameter	DRC-SFM Post. Mean	DRC-SFM Post. Std.	DRC-E-SFM Post. Mean	DRC-E-SFM Post. Std.
Frontier Parameters					
$const$	$\bar{\alpha}$	0.861	0.046	0.966	0.062
y_{it-1}	$\bar{\rho}$	0.214	0.018	0.168	0.015
k_{it}	$\bar{\beta}_k$	0.090	0.008	0.102	0.008
l_{it}	$\bar{\beta}_l$	0.073	0.009	0.086	0.011
m_{it}	$\bar{\beta}_m$	0.618	0.017	0.637	0.017
Trend	$\bar{\beta}_t$	-0.006	0.001	-0.005	0.001
Ln Unc. Index	$\bar{\beta}_{UI}$	-0.002	0.003	-0.003	0.004
Parameters for Instruments					
$const$	$\bar{\lambda}_0$			10.991	0.181
$PC1$	$\bar{\lambda}_1$			-0.003	0.003
$PC2$	$\bar{\lambda}_2$			-0.024	0.003
$PC3$	$\bar{\lambda}_3$			0.015	0.003
t	$\bar{\lambda}_4$			0.035	0.005
$Corr(v_{it}; e_{it})$	θ			0.070	0.034
e_{it} Std	σ_e			0.287	0.004
Returns to Scale					
RTS	$\bar{\beta}_k + \bar{\beta}_l + \bar{\beta}_m$	0.781	0.018	0.824	0.015
Estimated Technical Efficiencies					
Mean		0.901		0.889	
Median		0.931		0.933	
Max		0.972		0.987	

Notes: The table reports the posterior average and the standard deviation of the parameters. The MCMC algorithm is based on 5,000 iterations from which the first 1,500 samples are discarded. To save space, we do not report the posterior distributions of σ_{vi}'s and σ_{ui}'s in this table.

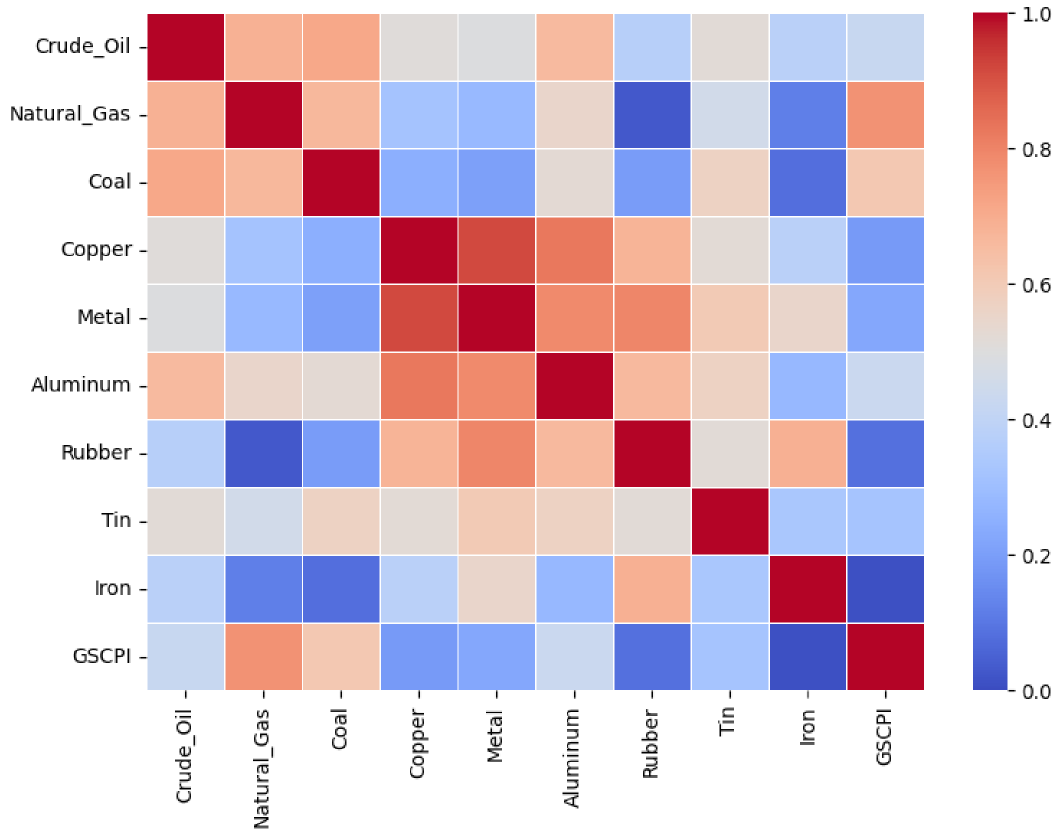

Figure 1: Correlation Heatmap of percentage changes (Y-o-Y)

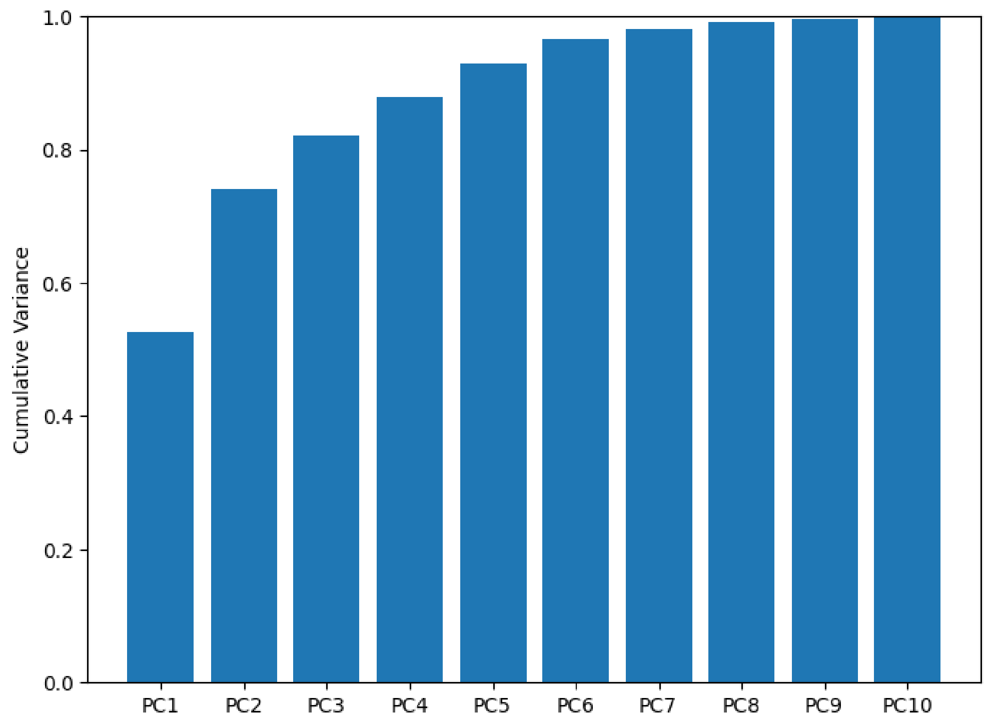

Figure 2: Cumulative Variance of PC Analysis

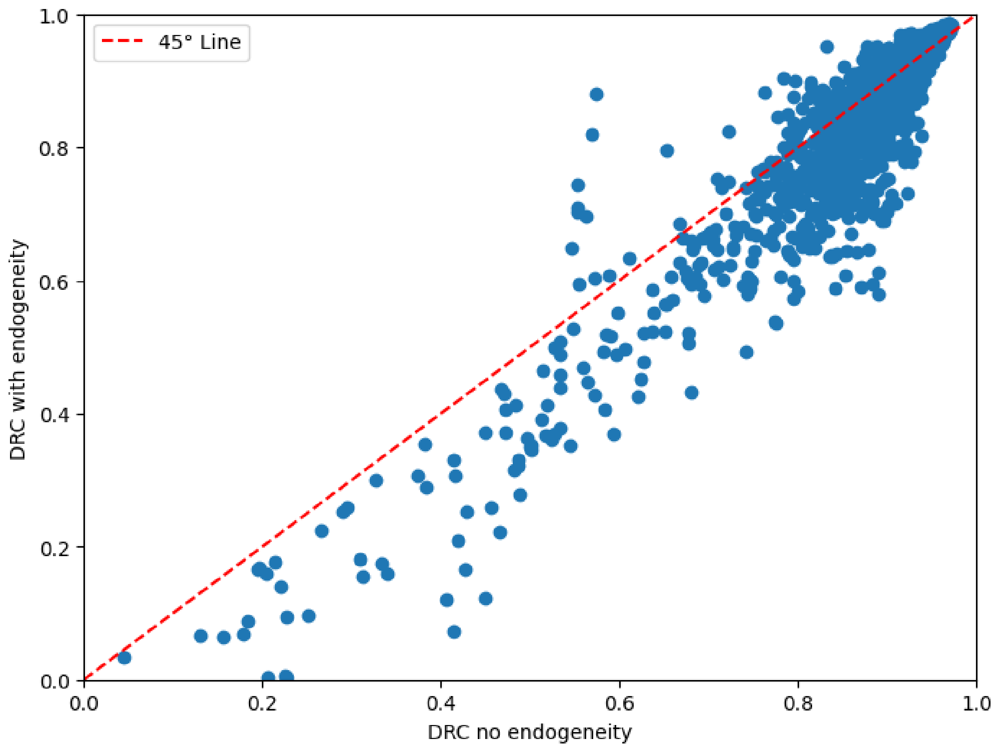

Figure 3: Efficiency Estimates for DRC-SFM and DRC-E-SFM

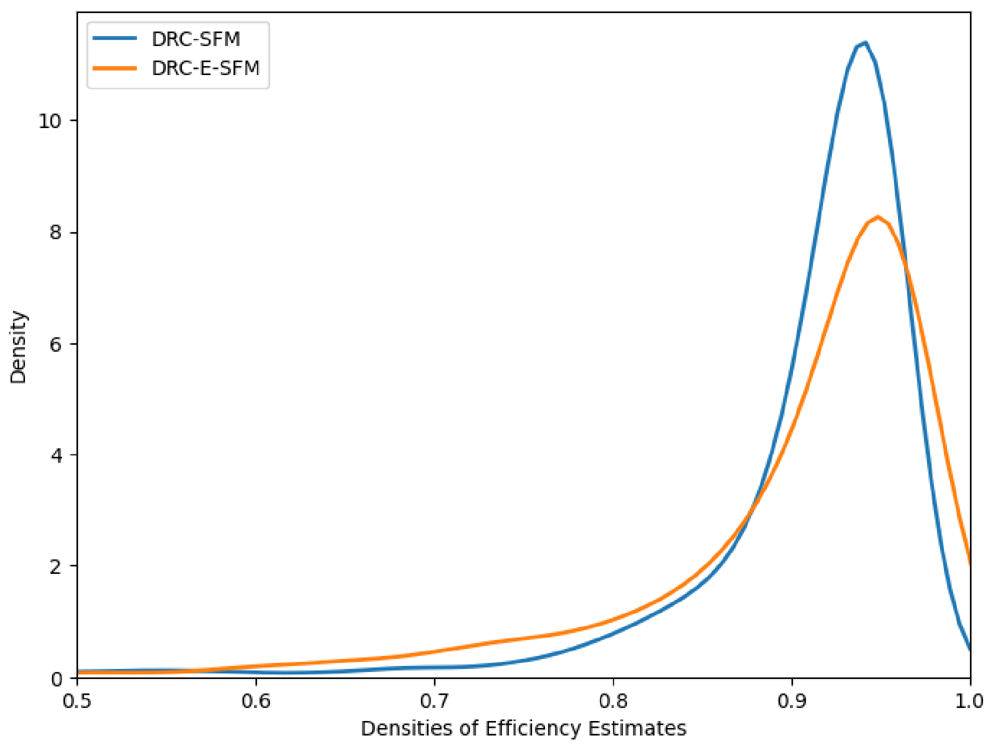

Figure 4: Kernel Densities of Efficiency Estimates for DRC-SFM and DRC-E-SFM

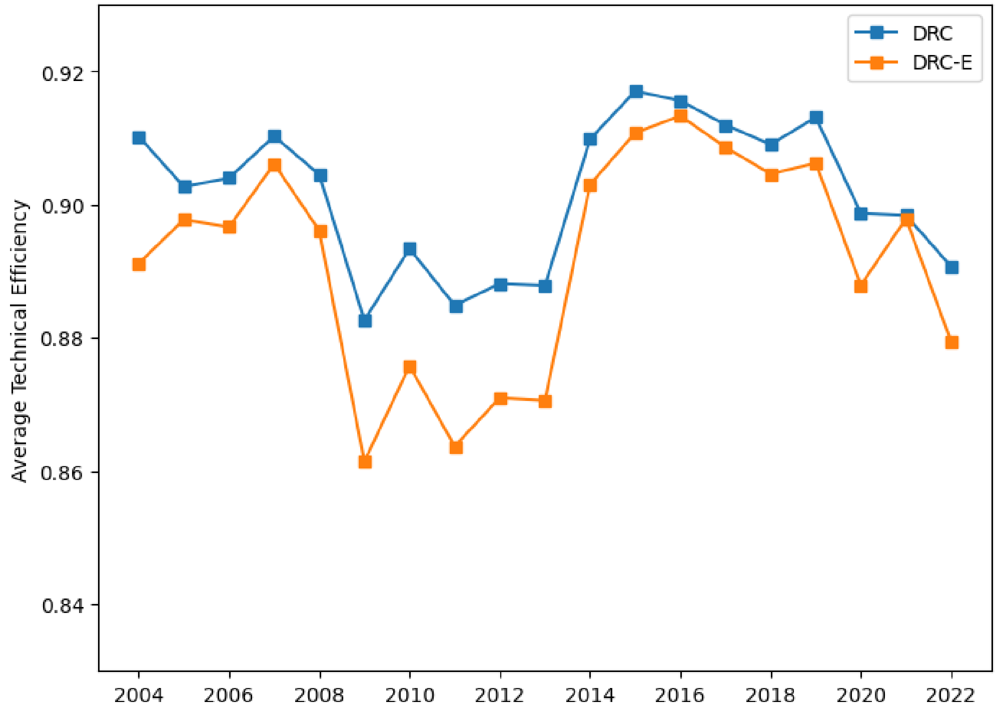

Figure 5: Intertemporal Average Efficiency for DRC-SFM and DRC-E-SFM

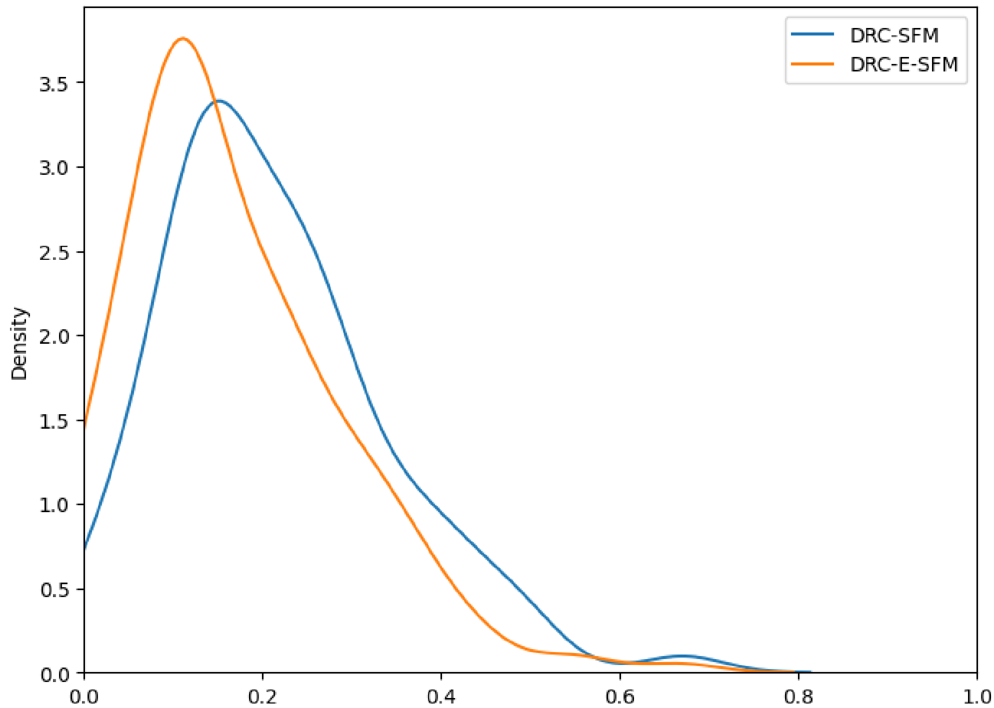

Figure 6: Kernel Densities of Autoregressive Parameter Estimates for DRC-SFM and DRC-E-SFM

Appendix A

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 augmented posterior distribution in (24). The iterated algorithm is used to approximate the marginal distribution of the corresponding parameters. To obtain a sample $\{\boldsymbol{\Theta}^s$ with $s = 1, 2, \ldots, S\}$ that converges to the marginal distribution of each unknown parameter, we can follow the procedure below:

- Step 1: Draw $\boldsymbol{\delta}_i \sim p(\boldsymbol{\delta}_i|\boldsymbol{\Theta}_{-\boldsymbol{\delta}_i})$
- Step 2: Draw $\bar{\boldsymbol{\delta}} \sim p(\bar{\boldsymbol{\delta}}|\boldsymbol{\Theta}_{-\bar{\boldsymbol{\delta}}})$
- Step 3: Draw $\boldsymbol{\lambda}_i \sim p(\boldsymbol{\delta}_i|\boldsymbol{\Theta}_{-\boldsymbol{\lambda}_i})$
- Step 4: Draw $\bar{\boldsymbol{\lambda}} \sim p(\bar{\boldsymbol{\delta}}|\boldsymbol{\Theta}_{-\bar{\boldsymbol{\lambda}}})$
- Step 5: Draw $\boldsymbol{\Omega} \sim p(\boldsymbol{\Omega}|\boldsymbol{\Theta}_{-\boldsymbol{\Omega}})$
- Step 6: Draw $\boldsymbol{A} \sim p(\boldsymbol{\Omega}|\boldsymbol{\Theta}_{-\boldsymbol{A}})$
- Step 7: Draw $\sigma_v^2 \sim p(\sigma_v^2|\boldsymbol{\Theta}_{-\sigma_v^2})$
- Step 8: Draw $\sigma_u^2 \sim p(\sigma_u^2|\boldsymbol{\Theta}_{-\sigma_u^2})$
- Step 9: Draw $\sigma_e^2 \sim p(\sigma_e^2|\boldsymbol{\Theta}_{-\sigma_e^2})$
- Step 10: Draw $\theta \sim p(\theta|\boldsymbol{\Theta}_{-\theta})$
- Step 11: Draw $\boldsymbol{u}_i \sim p(\boldsymbol{u}_i|\boldsymbol{\Theta}_{-\boldsymbol{u}_i})$ for $i = 1, 2, \ldots, N$
- Step 12: Draw $\boldsymbol{y}_0 \sim p(\boldsymbol{y}_0|\boldsymbol{\Theta}_{-\boldsymbol{y}_0})$
- Step 13: Repeat the above steps S times

Once we obtain the S sample from the conditional distributions of the parameter vector $\boldsymbol{\Theta}$, the posterior mean estimates can be obtained as:

$$E(\boldsymbol{\Theta}) = \int \boldsymbol{\Theta} p(\boldsymbol{\Theta}|.) \approx \frac{1}{S}\sum_{s=1}^{S} \boldsymbol{\Theta}^s$$

where $\boldsymbol{\Theta}^s$ is the s^{th} draw from the conditonal distribution.

Conditional Posterior Distributions

In this section we derive the conditional distributions for all parameters.

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

$$\boldsymbol{\delta}_i|\boldsymbol{\Theta}_{-\boldsymbol{\delta}_i} \sim \boldsymbol{\mathcal{N}}(\hat{\boldsymbol{b}}_i, \hat{\boldsymbol{V}}_i) \tag{28}$$

where

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

and

$$\tilde{\boldsymbol{y}}_i = \boldsymbol{y}_i - \theta\frac{\sigma_v}{\sigma_e}(\boldsymbol{q}_i - \boldsymbol{Z}_i\boldsymbol{\lambda}_i)$$

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

$$\bar{\boldsymbol{\delta}}|\boldsymbol{\Theta}_{-\bar{\boldsymbol{\delta}}} \sim \boldsymbol{\mathcal{N}}(\boldsymbol{B}[N\boldsymbol{\Omega}^{-1}\tilde{\boldsymbol{\delta}} + \boldsymbol{\Lambda}_0^{-1}\boldsymbol{\mu}_0], \boldsymbol{B}) \tag{29}$$

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 $\boldsymbol{\lambda}_i$'s:**

To derive the conditional densities of $\boldsymbol{\lambda}_i$'s, first we notice that $\boldsymbol{\lambda}_i$ is involved in the following equations:

$$\tilde{\boldsymbol{q}}_i = \boldsymbol{Z}_i\boldsymbol{\lambda}_i + \boldsymbol{\chi}_i$$
$$\boldsymbol{q}_i = \boldsymbol{Z}_i\boldsymbol{\lambda}_i + \boldsymbol{e}_i$$
$$\bar{\boldsymbol{\lambda}} = \boldsymbol{I}_{J_2\times J_2}\boldsymbol{\lambda}_i + \boldsymbol{\zeta}_i$$

where $\boldsymbol{\chi}_i \sim \mathcal{N}\left(\boldsymbol{0}, \sigma_e^2\frac{1-\theta^2}{\theta^2}\boldsymbol{I}_{T\times T}\right)$, $\boldsymbol{e}_i \sim \mathcal{N}\left(\boldsymbol{0}, \sigma_e^2\boldsymbol{I}_{T\times T}\right)$, $\boldsymbol{\zeta}_i \sim \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{A}\right)$ and:

$$\tilde{\boldsymbol{q}}_i \equiv \boldsymbol{q}_i - (\boldsymbol{y}_i - \boldsymbol{W}_i\boldsymbol{\delta}_i + \boldsymbol{u}_i)\frac{\sigma_e}{\theta\sigma_v}$$

Thus, we can show:

$$\boldsymbol{\lambda}_i|\boldsymbol{\Theta}_{-\boldsymbol{\lambda}_i} \sim \boldsymbol{\mathcal{N}}(\hat{\boldsymbol{\lambda}}_i, \hat{\boldsymbol{\Xi}}_i) \tag{30}$$

where

$$\hat{\boldsymbol{\lambda}}_i = \left(\boldsymbol{Z}_i'\boldsymbol{Z}_i\left(\frac{1}{\sigma_e^2\frac{1-\theta^2}{\theta^2}} + \frac{1}{\sigma_e^2}\right) + \boldsymbol{A}^{-1}\right)^{-1}\left(\frac{\boldsymbol{Z}_i'\tilde{\boldsymbol{q}}_i}{\sigma_e^2\frac{1-\theta^2}{\theta^2}} + \frac{\boldsymbol{Z}_i'\boldsymbol{q}_i}{\sigma_e^2} + \boldsymbol{A}^{-1}\bar{\boldsymbol{\lambda}}\right)$$

$$\hat{\boldsymbol{\Xi}}_i = \left(\boldsymbol{Z}_i'\boldsymbol{Z}_i\left(\frac{1}{\sigma_e^2\frac{1-\theta^2}{\theta^2}} + \frac{1}{\sigma_e^2}\right) + \boldsymbol{A}^{-1}\right)^{-1}$$

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

$$\bar{\boldsymbol{\lambda}}|\boldsymbol{\Theta}_{-\bar{\boldsymbol{\lambda}}} \sim \boldsymbol{\mathcal{N}}(\boldsymbol{B}[N\boldsymbol{A}^{-1}\tilde{\boldsymbol{\lambda}} + \boldsymbol{\Gamma}_0^{-1}\boldsymbol{\lambda}_0], \boldsymbol{B}) \tag{31}$$

where

$$\tilde{\boldsymbol{\lambda}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\lambda}_i$$

$$\boldsymbol{B} = [N\boldsymbol{A}^{-1} + \boldsymbol{\Gamma}_0^{-1}]^{-1}$$

- **Conditional Distribution of $\boldsymbol{\Omega}$:**

$$\boldsymbol{\Omega}|\boldsymbol{\Theta}_{-\boldsymbol{\Omega}} \sim \boldsymbol{\mathcal{IW}}\left(N + v_0, \sum_{i=1}^{N}(\boldsymbol{\delta}_i - \bar{\boldsymbol{\delta}})(\boldsymbol{\delta}_i - \bar{\boldsymbol{\delta}})' + \boldsymbol{\Psi}_0\right) \tag{32}$$

- **Conditional Distribution of $\boldsymbol{A}$:**

$$\boldsymbol{A}|\boldsymbol{\Theta}_{-\boldsymbol{A}} \sim \boldsymbol{\mathcal{IW}}\left(N + v_0, \sum_{i=1}^{N}(\boldsymbol{\lambda}_i - \bar{\boldsymbol{\lambda}})(\boldsymbol{\lambda}_i - \bar{\boldsymbol{\lambda}})' + \boldsymbol{\Phi}_0\right) \tag{33}$$

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

$$\boldsymbol{u}_i|\boldsymbol{\Theta}_{-\boldsymbol{u}_i} \sim \boldsymbol{\mathcal{N}}^{+}\left(-\frac{\sigma_u^2(\tilde{\boldsymbol{y}}_i - \boldsymbol{W}_i\boldsymbol{\delta}_i)}{\sigma_u^2 + \sigma_v^2(1-\theta^2)}, \frac{\sigma_v^2(1-\theta^2)\sigma_u^2}{\sigma_u^2 + \sigma_v^2(1-\theta^2)}\right) \tag{34}$$

where

$$\tilde{\boldsymbol{y}}_i = \boldsymbol{y}_i - \theta\frac{\sigma_v}{\sigma_e}(\boldsymbol{q}_i - \boldsymbol{Z}_i\boldsymbol{\lambda}_i)$$

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

$$\sigma_u^2|\boldsymbol{\Theta}_{-\sigma_u^2} \sim \mathcal{IG}\left(\frac{NT}{2} + \gamma_0, \frac{\sum_{i=1}^{N}\boldsymbol{u}_i'\boldsymbol{u}_i}{2} + \gamma_1\right) \tag{35}$$

where γ_0 and γ_1 are the shape and scale parameters of the prior pdf, respectively.

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

$$f(\sigma_v^2|\boldsymbol{\Theta}_{-\sigma_v^2}) \propto (\sigma_v^2)^{-\frac{NT}{2}} exp\left\{-\frac{1}{2\sigma_v^2(1-\theta^2)}\sum_{i=1}^{N}\boldsymbol{\omega_i'\omega_i}\right\}\pi(\sigma_v^2|a_0, a_1)$$

where $\boldsymbol{\omega_i} = \boldsymbol{y}_i - \boldsymbol{W}_i\boldsymbol{\delta}_i - \theta\frac{\sigma_v}{\sigma_e}(\boldsymbol{q}_i - \boldsymbol{Z}_i\boldsymbol{\lambda}_i) + \boldsymbol{u}_i$, and $\pi(\sigma_v^2|a_0, a_1)$ is the Inverse Gamma prior distribution. The above kernel function does not belong to a known distributional form, and thus, sampling from it is not straightforward. Therefore, we use a Metropolis Hastings algorithm where we generate a candidate σ_v^{cand} from a proposal distribution $q(\sigma_v^{cand}|.)$, and we accept the candidate parameter using a probability. The probability is given by:

$$p = min\left\{1, \frac{f(\sigma_v^{cand}|.)}{f(\sigma_v^c|.)}\frac{q(\sigma_v^c|.)}{q(\sigma_v^{cand}|.)}\right\}$$

where σ_v^c is the previously accepted parameter draw.

Here we note that when $\theta = 0$, we arrive at:

$$\sigma_v^2|\boldsymbol{\Theta}_{-\sigma_v^2} \sim \mathcal{IG}\left(\frac{NT}{2} + a_0, \frac{\sum_{i=1}^N(\boldsymbol{y}_i - \boldsymbol{W}_i\boldsymbol{\delta}_i + \boldsymbol{u}_i)'(\boldsymbol{y}_i - \boldsymbol{W}_i\boldsymbol{\delta}_i + \boldsymbol{u}_i)}{2} + a_1\right) \quad (36)$$

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

$$f(\sigma_e^2|\boldsymbol{\Theta}_{-\sigma_e^2}) \propto (\sigma_e^2)^{-\frac{NT}{2}} exp\left\{-\frac{1}{2\sigma_v^2(1-\theta^2)}\sum_{i=1}^N \boldsymbol{\omega_i'\omega_i}\right\} \quad (37)$$
$$exp\left\{-\frac{1}{2\sigma_e^2}\sum_{i=1}^N(\boldsymbol{q}_i - \boldsymbol{Z}_i\boldsymbol{\lambda}_i)'(\boldsymbol{q}_i - \boldsymbol{Z}_i\boldsymbol{\lambda}_i)\right\}\pi(\sigma_e^2|h_0, h_1)$$

where $\boldsymbol{\omega_i} = \boldsymbol{y}_i - \boldsymbol{W}_i\boldsymbol{\delta}_i - \theta\frac{\sigma_v}{\sigma_e}(\boldsymbol{q}_i - \boldsymbol{Z}_i\boldsymbol{\lambda}_i) + \boldsymbol{u}_i$, and $\pi(\sigma_e^2|h_0, h_1)$ is the Inverse Gamma prior distribution. The above kernel function does not belong to a known distributional form. We follow, a Metropolis Hastings algorithm where we generate a candidate σ_e^{cand} from a proposal distribution $p(\sigma_e^{cand}|.)$, and we accept the candidate parameter using a probability. The probability is given by:

$$p = min\left\{1, \frac{f(\sigma_e^{cand}|.)}{f(\sigma_e^c|.)}\frac{q(\sigma_e^c|.)}{q(\sigma_e^{cand}|.)}\right\}$$

where σ_e^c is the previously accepted parameter draw.

- **Conditional distribution of θ:**

$$f(\theta|\boldsymbol{\Theta}_{-\theta}) \propto (1-\theta^2)^{-\frac{NT}{2}} exp\left\{-\frac{1}{2\sigma_v^2(1-\theta^2)}\sum_{i=1}^N \boldsymbol{\omega_i'\omega_i}\right\}\pi(\theta|b_0, b_1) \quad (38)$$

where $\boldsymbol{\omega_i} = \boldsymbol{y}_i - \boldsymbol{W}_i\boldsymbol{\delta}_i - \theta\frac{\sigma_v}{\sigma_e}(\boldsymbol{q}_i - \boldsymbol{Z}_i\boldsymbol{\lambda}_i) + \boldsymbol{u}_i$, and $\pi(\theta|b_0, b_1)$ is the Beta prior distribution. We follow, a Metropolis Hastings algorithm where we generate a candidate θ^{cand} from a proposal distribution $p(\theta^{cand}|.)$, and we accept the candidate parameter using a probability. The probability is given by:

$$p = min\left\{1, \frac{f(\theta^{cand}|.)}{f(\theta^c|.)}\frac{q(\theta^c|.)}{q(\theta^{cand}|.)}\right\}$$

where θ^c is the previously accepted parameter draw.

- **Conditional Distribution of y_{i0}'s:**

$$y_{i0}|\boldsymbol{\Theta}_{-y_{i0}} \sim \mathcal{N}\left(\frac{\rho_i\sigma_0^2\omega_{i1} + \sigma_v^2(1-\theta^2)\bar{y}_0}{\sigma_0^2\rho_i^2 + \sigma_v^2(1-\theta^2)}, \frac{\sigma_v^2(1-\theta^2)\sigma_0^2}{\sigma_0^2\rho_i^2 + \sigma_v^2(1-\theta^2)}\right) \tag{39}$$

where

$$\omega_{i1} = y_{i1} - \boldsymbol{w}'_{i1}\boldsymbol{\delta}_i - \theta\frac{\sigma_v}{\sigma_e}(q_{i1} - \boldsymbol{z}'_{i1}\boldsymbol{\delta}_i) + u_{i1}$$

Efficiency Measurement

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

$$\boldsymbol{u}_i^1, \boldsymbol{u}_i^2, \boldsymbol{u}_i^3, ..., \boldsymbol{u}_i^S \tag{40}$$

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:

$$\hat{\boldsymbol{u}}_i = S^{-1}\sum_{s=1}^{S}\boldsymbol{u}_i^s \tag{41}$$

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

$$\hat{\boldsymbol{TE}}_i = exp(-\hat{\boldsymbol{u}}_i) \tag{42}$$

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

Appendix B

Table 9: UK 2-digit SIC (2007) Codes

Manufacturing Sector		Construction Sector	
Code	Description	Code	Description
10	Manufacture of food products	41	Construction of buildings
11	Manufacture of beverages	42	Civil engineering
12	Manufacture of tobacco products	43	Specialised construction activities
13	Manufacture of textiles		
14	Manufacture of wearing apparel		
15	Manufacture of leather and related products		
16	Manufacture of wood and of products of wood and cork		
17	Manufacture of paper and paper products		
18	Printing and reproduction of recorded media		
19	Manufacture of coke and refined petroleum products		
20	Manufacture of chemicals and chemical products		
21	Manufacture of basic pharmaceutical products		
22	Manufacture of rubber and plastic products		
23	Manufacture of other non-metallic mineral products		
24	Manufacture of basic metals		
25	Manufacture of fabricated metal products		
26	Manufacture of computer; electronic and optical products		
27	Manufacture of electrical equipment		
28	Manufacture of machinery and equipment		
29	Manufacture of motor vehicles		
30	Manufacture of other transport equipment		
31	Manufacture of furniture		
32	Other manufacturing		
33	Repair and installation of machinery and equipment		

Appendix C

In this appendix, we evaluate the relevance of our instruments. We follow Tsionas et al. (2023) and we compute the generalized R-squared R^2_*, to evaluate how well our exogenous instruments can proxy the flexible input m_{it}. Specifically, we focus on our equation:

$$m_{it} = \lambda_{0i} + \lambda_{1i} PC1_{it} + \lambda_{2i} PC2_{it} + \lambda_{3i} PC3_{it} + \lambda_{4i} t + e_{it}$$

It is clear, that if all instruments are irrelevant, all parameters would be zero (except the constant), and thus the equation above would be able to capture only the average effect for each $i = 1, 2, \dots, N$. In this scenario, the standard deviation of e_{it} would be $\sigma_{ir} = 0.446$. Therefore, the generalized R^2_* can be computed as:

$$R^2_* = 1 - \frac{\sigma_e^2}{\sigma_{ir}^2}$$

For each MCMC draw, we obtain the σ_e and compute the generalized R^2_{*s}. Figure 7 illustrates the density of R^2_*. It is clear that the proposed instruments are capable of explaining more than 55% of the total variation of the endogenous input m_{it}. To further assess the instruments,

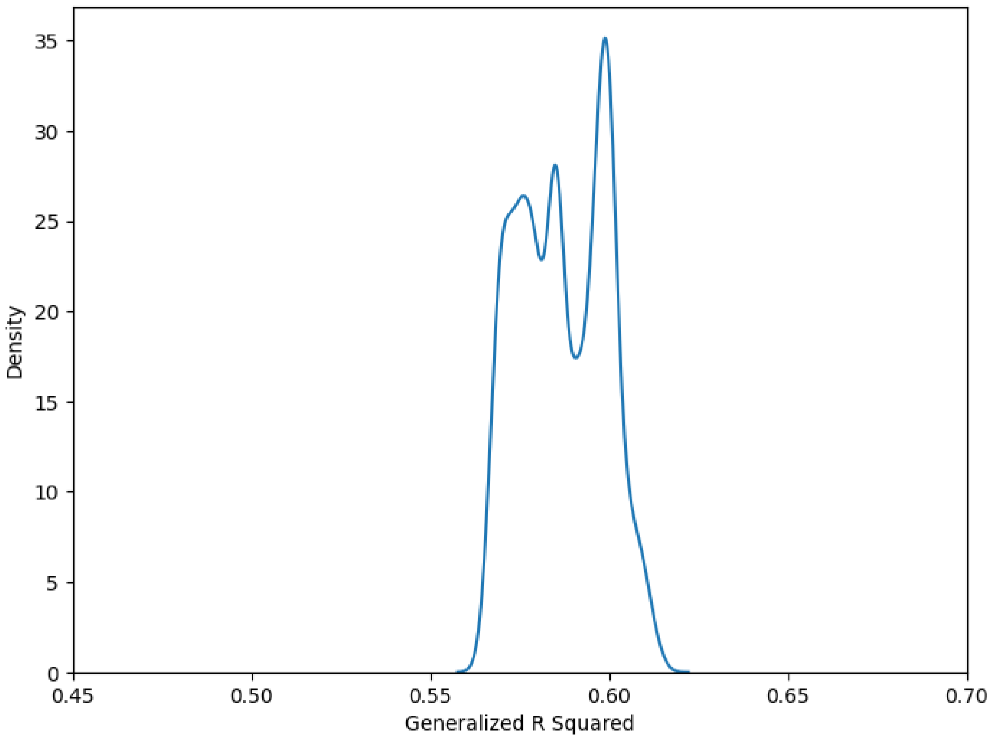

Figure 7: Density of estimated generalized R-squared

we follow the vast majority of weak instruments literature, where researchers compute the F-statistic of the first stage and compare it with empirical rule of thumb cut-off of 10! Thus, in our setting, for each MCMC iteration, we compute a pseudo F-statistic given by:

$$F_* = \frac{R_*^2/(NT-k)}{(1-R_*^2)/(NT-k-1)}$$

where NT is total number of observations and k is the number of variables used to approximate m_{it}. Figure 8 illustrates the density of F_* obtained from the MCMC iterations. It is evident that all F_*s are much higher than the empirical cut-off value.

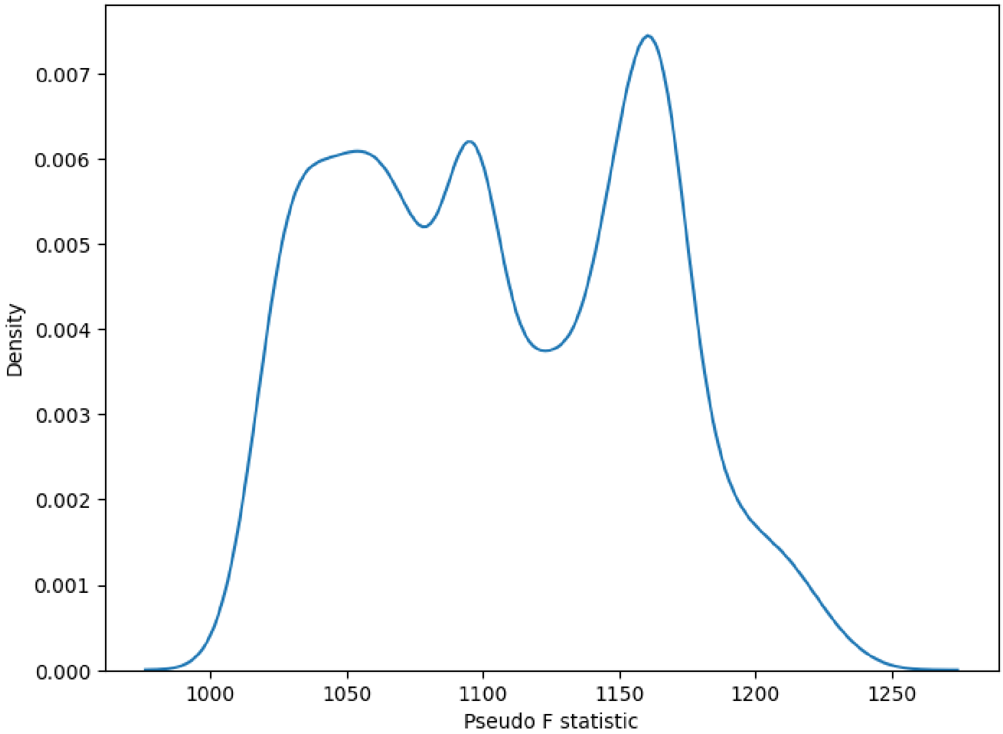

Figure 8: Density of estimated pseudo F-statistic

Appendix D

In this appendix, we report the sensitive analysis of our results for different inefficiency prior specifications. Given that $\sigma_{ui} \sim \mathcal{IG}(\gamma_0, \gamma_1)$ for all $i = 1, 2, \ldots, N$, we follow Van den Broeck et al. (1994) and set different values for the shape γ_0 and scale γ_1 parameters. Specifically, given a prior median efficiency level r^*, we can set $\gamma_0 = 5$ and $\gamma_1 = 10ln(r^*)^2$.

We re-estimate the DRC-SFM and DRC-E-SFM models for the following prior median efficiency beliefs:

$$r^* = \{0.5, 0.55, 0.6, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95\}$$

We note that for the all the rest parameters, we use the standard non-informative priors, as presented in the main text. All models have been re-estimated using 5,000 draws from which the first 1,500 have been discarded to eliminate any initial value effects.

For the different model estimates we save posterior averages of the $\bar{\rho}$, θ, short-run RTS (SR-RTS), long-run LR-RTS (LR-RTS), as well as the average and median efficiency estimates. We believe that lower r^* values consist of extreme scenarios and reflect unrealistic prior beliefs, since publicly traded manufacturing firms cannot operate at low efficiency levels in such a competitive global environment. Table 7 reports the posterior averages for the different r^*, as well as the estimates under the non-informative prior (NI) presented in the main body. Overall, we see that lower prior median efficiency beliefs generate higher autoregressive parameter estimates and lower SR-RTS. In addition, we see that under all the various specifications, the LR-RTS estimates are very close to one, implying that firms operate under flat average costs, ceteris paribus.

Last, we comment on the median efficiency estimates for the DRC-SFM and DRC-E-SFM models, respectively. Overall, we find that there is a possible correlation between the prior median efficiency and the posterior median efficiency estimate. This is expected as setting the prior median score to a specific value could generate an informative prior. However, it is clear for both models, there is a lower and upper bound for those estimates, which can be seen as their extreme upper and lower estimates. Specifically, we see that for the DRC-SFM, the lower and the higher median efficiency levels are 75.9% and 94.7%, respectively. Similarly, the DRC-E-SFM produces median efficiency scores of 79% and 94.6% for the lower and upper bounds, respectively.

Table 10: Posterior Parameter Estimates for different Prior Median Efficiency Beliefs

	Prior Median Efficiency Belief r^*										
	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	NI
	DRC-SFM										
$\bar{\rho}$	0.347	0.343	0.328	0.309	0.300	0.285	0.268	0.250	0.236	0.214	**0.214**
SR-RTS	0.650	0.651	0.667	0.687	0.696	0.712	0.731	0.752	0.775	0.794	**0.781**
LR-RTS	0.995	0.991	0.993	0.994	0.994	0.996	0.999	1.003	1.014	1.010	**0.994**
Mean Eff.	0.742	0.759	0.776	0.794	0.812	0.831	0.853	0.877	0.906	0.945	**0.901**
Median Eff.	0.759	0.776	0.794	0.813	0.831	0.850	0.870	0.891	0.915	0.947	**0.931**
	DRC-E-SFM										
$\bar{\rho}$	0.315	0.305	0.288	0.264	0.246	0.235	0.219	0.201	0.185	0.162	**0.168**
RTS	0.673	0.690	0.705	0.726	0.751	0.758	0.766	0.794	0.814	0.838	**0.824**
θ	0.053	0.065	0.048	0.057	0.051	0.056	0.050	0.045	0.037	0.064	**0.070**
LR-RTS	0.990	0.992	0.990	0.986	0.996	0.990	0.980	0.993	1.015	1.000	**0.990**
Mean Eff.	0.769	0.780	0.794	0.812	0.824	0.838	0.854	0.868	0.885	0.908	**0.889**
Median Eff	0.790	0.800	0.817	0.837	0.849	0.868	0.884	0.902	0.920	0.946	**0.933**