Endogenous Dynamic Efficiency in the Intertemporal Optimization Models of Firm Behavior

Tsionas, Mike G. and Malikov, Emir and Kumbhakar, Subal C.

Lancaster University Management School, University of Nevada, Las Vegas, State University of New York at Binghamton

27 November 2019

Endogenous Dynamic Efficiency in the Intertemporal Optimization Models of Firm Behavior*

Mike G. Tsionas[1] Emir Malikov[2] Subal C. Kumbhakar[3,4]

[1]Lancaster University Management School, Lancaster, United Kingdom
[2]University of Nevada, Las Vegas, NV, United States
[3]State University of New York at Binghamton, Binghamton, NY, United States
[4]University of Stavanger Business School, Stavanger, Norway

November 27, 2019

Abstract

Existing methods for the measurement of technical efficiency in the dynamic production models obtain it from the implied distance functions without making use of the information about intertemporal economic behavior in the estimation beyond an indirect appeal to duality. The main limitation of such an estimation approach is that it does not allow for the dynamic evolution of efficiency that is explicitly optimized by the firm. This paper introduces a new conceptualization of efficiency that directly enters the firm's intertemporal production decisions and is both explicitly costly and endogenously determined. We build a moment-based multiple-equation system estimation procedure that incorporates both the dynamic and static optimality conditions derived from the firm's intertemporal expected cost minimization. We operationalize our methodology using a modified version of a Bayesian Exponentially Tilted Empirical Likelihood adjusted for the presence of dynamic latent variables in the model, which we showcase using the 1960–2004 U.S. agricultural farm production data. We find that allowing for potential endogenous adjustments in efficiency over time produces significantly higher estimates of technical efficiency, which is likely due to inherent inability of the more standard exogenous-efficiency model to properly credit firms for incurring efficiency-improvement adjustment costs. Our results also suggest material improvements in efficiency over time at an about 2.6% average annual rate, which contrasts with near-zero estimates of the exogenous efficiency change.

Keywords: dynamic efficiency, endogenous efficiency, intertemporal optimization, Bayesian analysis, Markov Chain Monte Carlo, sequential Monte Carlo

JEL Classification: C11, D24, Q10

* *Email:* m.tsionas@lancaster.ac.uk (Tsionas), emir.malikov@unlv.edu (Malikov), kkar@binghamton.edu (Kumbhakar).

We would like to thank the editor, Robert Dyson, and three anonymous referees for many insightful comments that helped improve this paper. Any remaining errors are our own.

1 Introduction

The use of explicit dynamic models of the firm's intertemporal production decisions to account for and provide estimates of technical (in)efficiency are still relatively scarce in the efficiency analysis literature. Most existing studies rely on the static formulations of firm production (e.g., see Parmeter & Kumbhakar, 2014; Kumbhakar, Parmeter & Zelenyuk, 2017, for recent surveys). However, for such static production models to provide a reasonably accurate description of the firm behavior, the firm ought to be able to freely adjust all its inputs at any point in time which would obviate any dynamic implications of its production decisions. In practice though, not all factors of production are freely-varying in the face of adjustment frictions (e.g., time-to-install, hiring costs) that may render them quasi-fixed in the short run. This notable weakness of the static production models (and hence of the static inefficiency measurements derived therefrom) owing to their inability to accommodate gradual adjustments in dynamic inputs, whereby the present production decisions have dynamic implications for future production outcomes, has recently given rise to a new strand of literature focused on incorporating intertemporal aspects of firm optimization into the efficiency measurement. The work of Silva & Stefanou (2003, 2007) is a fundamental advance on these issues, wherein the authors develop a dynamic optimization model of the firm production behavior using intertemporal cost minimization with quasi-fixed factors, based upon which they propose a nonparametric[1] measure of dynamic (technical) efficiency (see Kapelko & Oude Lansink, 2017, for a multi-direction extension). Alternatively, Rungsuriyawiboon & Stefanou (2007) use a shadow cost approach in a similar intertemporal cost minimization framework to nonparametrically recover efficiency measures under dynamic duality. For some earlier but simpler attempts to nonparametrically model dynamic aspects of the firm production, also see Sengupta (1995) and Nemoto & Goto (1999, 2003). More recently, Serra, Oude Lansink & Stefanou (2011) and Silva & Oude Lansink (2013) have proposed a measurement of dynamic efficiency based on a primal directional-distance-function-based representation of production technology which has been gaining popularity in the literature partly because it can be operationalized via both the data envelopment and (econometric) stochastic frontier methods.[2] This formulation has since been extended to enable a decomposition of dynamic efficiency (Kapelko, Oude Lansink & Stefanou, 2014) and to also measure "dynamic" productivity growth (Oude Lansink, Stefanou & Serra, 2015), with multiple empirical applications that have followed.

While the theoretical models of dynamic production decisions do explicitly postulate the firm's intertemporal optimization subject to adjustment frictions in quasi-fixed inputs, available methodologies for the estimation of dynamic efficiency in such frameworks in practice hardly make use of the information embedded in such dynamic optimization problems. Rather, the popular go-to approach

[1] Estimated using the (non-stochastic) data envelopment analysis.

[2] The closely related literature includes studies of dynamic production systems in the presence of intertemporal inefficiency and the lagged effects (Chen & van Dalen, 2010; Cherchye, De Rock & Kerstens, 2018), studies of inefficiency under sequential technologies (Hampf, 2017) as well as the work that seeks to dynamize evolution of time-persistent inefficiency in a more reduced dynamic framework (Wang & Huang, 2007; Skevas, Emvalomatis & Brümmer, 2018).

is to estimate technical efficiency from Silva & Oude Lansink's (2013) dynamic directional distance function (oriented in the space of freely-varying inputs and investments into quasi-fixed dynamic factors) under its duality to the optimal current value function associated with the firm's intertemporal cost-minimization problem under the Hamilton-Jacobi-Bellman conditions (e.g., Serra et al., 2011; Kapelko et al., 2014; Oude Lansink et al., 2015; Kapelko, Oude Lansink & Stefanou, 2016, 2017; Minviel & Sipiläinen, 2018). Thus, in practice, the strategy for a (primal) estimation of dynamic technical efficiency is effectively the same as that for the estimation of a more conventional static efficiency except that in the former case a distance to the frontier is measured in the space of variable inputs *and investments* as opposed to merely conditioning the distance function on quasi-fixed levels of dynamic inputs. For instance, if one were to adopt econometric techniques, the estimation then boils down to a familiar procedure of appending a convoluted stochastic error containing a one-sided latent inefficiency and a two-sided noise to the distance function *a posteriori*. Such an estimation approach may however be overly restrictive because it (*i*) does not explicitly account for dynamics in efficiency itself as well as does not allow for the costly firm-controlled efficiency change, (*ii*) neglects the likely possibility that past dynamic efficiency is a part of the information set based upon which the firm optimizes, (*iii*) makes no use of information about economic behavior in the estimation beyond an indirect appeal to duality despite seeking a deeply structural "dynamic" interpretation of efficiency, and (*iv*) suffers from the endogeneity problem due to simultaneity of the variable input and investment decisions. Lastly and perhaps more importantly, the underlying conceptual framework of dynamic production that the available estimation methodologies are based upon does not allow for the dynamic evolution of efficiency that is explicitly optimized by the firm. While conditioning of the distance function on investments in dynamic inputs implicitly allows a departure from the strong assumption of exogenous efficiency by letting the firm to indirectly control the evolution of its efficiency (through investment decisions), such an endogenization of dynamic efficiency is derivative/indirect in its nature. It also implicitly assumes away the costliness of efficiency changes beyond the costs derived from frictions in dynamic inputs thereby dispensing with the expenses pertaining specifically to efficiency improvements such as adoption of quality control and other improved business practices, management training, etc.

In this paper, we contribute to the literature by developing a new (structural) conceptualization of dynamic efficiency that directly enters the firm's intertemporal production decisions and is both explicitly costly and endogenously determined. We differentiate between the variable-inputs-oriented inefficiency and factor-specific distortions in quasi-fixed inputs and allow all of these dynamic latent variables to jointly evolve over time. We then build a moment-based multiple-equation system estimation procedure that incorporates the variable cost function and both the dynamic and static optimality conditions derived from the firm's intertemporal minimization of a discounted stream of future costs. Not only does this system approach let us handle endogeneity of inputs (and, implicitly, investments too), but it also facilitates a structurally meaningful interpretation of the endogenous dynamic technical efficiency that evolves following its optimal path consistent with the firm's intertemporal cost-minimizing objective.

Our methodology for the estimation of dynamic efficiency therefore provides a more elaborate alternative to that based on dynamic distance functions estimated under duality which restrictively treats efficiency as exogenous and distributed independently over time (e.g., see Serra et al., 2011). Now, ours is not the first attempt to explicitly model temporal dynamics of latent technical efficiency econometrically, where a few earlier works include Ahn, Good & Sickles (2000) and Tsionas (2006). However, to our knowledge, no prior study has done so in conjunction with a full intertemporal optimization problem of the firm like we do in this paper which, among other things, enables us to endogenize inefficiency as well as to explicitly accommodate its costliness in the firm optimization.[3] The latter is particularly desirable given the interest of economists in linking dynamic efficiency to adjustment costs and the real option values of investment (e.g., Lambarraa, Stefanou & Gil, 2016).

We showcase our model by applying it to an annual state-level panel data on agricultural farm production in lower 48 contiguous states of the United States during the 1960–2004 period. We follow Gallant, Giocomini & Ragusa (2017) to estimate models involving moment conditions and latent dynamic variables, although we do not use a Metropolis-Hastings approach or sequential importance sampling because Sequential Monte Carlo is more efficient. We use a modified version of a Bayesian Exponentially Tilted Empirical Likelihood adjusted for the presence of dynamic latent variables in the model (i.e., variable-inputs-oriented technical efficiency and quasi-fixed factor distortions) that, among other things, does not rely on using a fully parametric likelihood. The joint dynamics of latent variables is formulated using a second-order vector autoregression, with the choice of order motivated by temporal dynamics in the firm's Euler equations. Among other things, we find that the failure to allow for potential endogenous adjustments in efficiency over time produces significantly lower estimates of dynamic efficiency, which is likely due to the inherent inability of a more traditional exogenous-efficiency framework to properly credit producing units for incurring efficiency-improvement adjustment costs. The data overwhelmingly support our approach.

The rest of the paper proceeds as follows. Section 2 introduces our model of dynamic production decisions in the presence of endogenous efficiency. We describe the estimation details in Section 3. Section 4 reports the empirical application. We conclude in Section 5.

2 Model

Consider the dynamic production process. Let $\boldsymbol{x}_t \in \Re^J_+$ denote the vector of freely varying inputs with $\boldsymbol{w}_t \in \Re^J_{++}$ being the vector of corresponding prices. Further, let $\boldsymbol{k}_t \in \Re^M_+$ be the vector of quasi-fixed dynamic factors of production and $\boldsymbol{q}_t \in \Re^M_{++}$ denote their corresponding "rental" prices (user costs). Both $\boldsymbol{w}_t$ and $\boldsymbol{q}_t$ may evolve stochastically over time. Suppose also $\boldsymbol{\delta} = [\delta_1, \ldots, \delta_M]' \in [0,1]^M$ is a vector of depreciation rates, and $\beta \in [0, 1)$ is a time discount factor. The vector of outputs is given by $\boldsymbol{y}_t \in \Re^Q_+$.

In line with the tradition in the literature on dynamic efficiency, we consider an intertemporal

[3] Also, see Tsionas & Izzeldin (2018) for a *static* production model with costly inefficiency.

cost minimization but, unlike previous studies, incorporate inefficiency directly into the optimization problem as well as allow for uncertainty. The uncertainty, in this case, can arise from the lack of perfect foresight about future market conditions including the input prices that may evolve over time stochastically. We also assume firms are risk-neutral. The risk-neutrality assumption is made primarily so that, in the firm optimization problem under uncertainty, we can work directly with the expectation of a future stream of costs (the outcomes) without needing to consider the optimization of their expected utilities. In this choice, we have opted to follow the convention in the literature on costly investments into dynamic factors under uncertainty. More specifically, we build on the seminal work by Pindyck & Rotemberg (1983). While we recognize that risk-neutrality may be too restrictive of an assumption, the advantage of such a formulation is that it provide us with the tractable way of modeling the firm's intertemporal behavior that gives rise to the Euler conditions *in expectation* which we then use as a basis to form the estimating moment conditions. Notably, our setup is already more flexible than the existing work on dynamic efficiency that assumes that the forward-looking firms have perfect foresight (e.g., Silva & Stefanou, 2003, 2007; Rungsuriyawiboon & Stefanou, 2007; Serra et al., 2011; Silva & Oude Lansink, 2013).

Firms are said to operate in perfectly competitive factor markets rendering present and future input prices exogenous. To highlight key features of our modeling approach, we begin with a simpler standard framework with no inefficiency, which we then augment step by step to incorporate *(i)* explicitly costly and *(ii)* endogenously determined inefficiency.

Optimization without Inefficiency. The firm's more traditional intertemporal expected cost optimization problem with respect to dynamic inputs is given by

$$\min_{\boldsymbol{k}_t} \ \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left\{ C(\boldsymbol{w}_t, \boldsymbol{k}_t, \boldsymbol{y}_t) + \sum_{m=1}^{M} \left[q_{m,t} k_{m,t} + G_m \big(\underbrace{k_{m,t} - (1-\delta_m) k_{m,t-1}}_{i_{m,t}} \big) \right] \right\}, \tag{2.1}$$

where $\mathbb{E}_0$ denotes expectation at time $t = 0$ conditional on the available information; each mth quasi-fixed dynamic input follows its respective law of motion, i.e.,

$$k_{m,t} = i_{m,t} + (1-\delta_m) k_{m,t-1} \quad \forall\ m,$$

with $i_{m,t}$ being the gross investment flow into this input's stock; $G_m(\cdot)$ is the function representing adjustment costs in k_m $\forall$ m; and $C(\boldsymbol{w}, \boldsymbol{k}, \boldsymbol{y})$: $\Re^J_{++} \times \Re^M_{+} \times \Re^Q_{+} \to \Re_{+}$ is the restricted (short-run) variable cost function optimized in the freely-varying static factors subject to the already predetermined quasi-fixed $\boldsymbol{k}_t$:

$$C(\boldsymbol{w}_t, \boldsymbol{k}_t, \boldsymbol{y}_t) \equiv \min_{\boldsymbol{x}_t} \ \boldsymbol{w}_t' \boldsymbol{x}_t : \ F(\boldsymbol{x}_t, \boldsymbol{k}_t, \boldsymbol{y}_t) = 1, \tag{2.2}$$

with $F(\boldsymbol{x}_t, \boldsymbol{k}_t, \boldsymbol{y}_t) = 1$ being the transformation function descriptive of the firm's production process.

Optimization Subject to Exogenous Inefficiency. To introduce inefficiency, we begin by considering the firm's static optimization problem where we introduce a possibility for the over/under-use of inputs à la Kumbhakar (1997), i.e.,

$$\min_{\boldsymbol{x}_t} \ \boldsymbol{w}_t'\boldsymbol{x}_t : \ F(\vartheta_t \boldsymbol{x}_t, \boldsymbol{\eta}_t \odot \boldsymbol{k}_t, \boldsymbol{y}_t) = 1, \tag{2.3}$$

with the scalar $\boldsymbol{x}_t$-oriented technical inefficiency $\vartheta_t \geq 1$ measuring over-use in all freely varying inputs and each element of $\boldsymbol{\eta}_t = [\eta_{1,t}, \ldots, \eta_{M,t}]' \in \Re^M_{++}$ representing a distortion in the corresponding quasi-fixed factor (i.e., the over- or under-use thereof). Note that, to capture likely heterogeneity in the degree of fixity across $\boldsymbol{k}_t$, the factor distortions $\{\eta_{m,t}\}$ are permitted to vary across individual quasi-fixed inputs. The above problem is equivalent to

$$C\left(\boldsymbol{w}_t, \widetilde{\boldsymbol{k}}, \boldsymbol{y}_t\right) \equiv \min_{\widetilde{\boldsymbol{x}}_t} \ \vartheta_t^{-1} \boldsymbol{w}_t' \widetilde{\boldsymbol{x}}_t : \ F\left(\widetilde{\boldsymbol{x}}_t, \widetilde{\boldsymbol{k}}_t, \boldsymbol{y}_t\right) = 1, \tag{2.4}$$

where $\widetilde{\boldsymbol{x}}_t \equiv \vartheta_t \boldsymbol{x}_t$ and $\widetilde{\boldsymbol{k}}_t \equiv \boldsymbol{\eta}_t \odot \boldsymbol{k}_t$ are the actual input quantities; and it is evident that the firm's actual cost is $C^a(\boldsymbol{w}_t, \widetilde{\boldsymbol{k}}_t, \boldsymbol{y}_t) = \vartheta_t C(\boldsymbol{w}_t, \widetilde{\boldsymbol{k}}_t, \boldsymbol{y}_t)$ and, hence, $\ln C^a(\boldsymbol{w}_t, \widetilde{\boldsymbol{k}}_t, \boldsymbol{y}_t) = \ln C(\boldsymbol{w}_t, \widetilde{\boldsymbol{k}}_t, \boldsymbol{y}_t) + u_t$ with $u_t \equiv \ln \vartheta_t \in \Re_+$ measuring variable cost inefficiency arising from the firm's technical inefficiency in the $\boldsymbol{x}_t$ orientation.

Thus, extending framework in (2.1)–(2.2) to explicitly allow for exogenous inefficiency and factor distortions taken by the firm as given, we obtain the following dynamic optimization problem:

$$\min_{\widetilde{\boldsymbol{k}}_t} \ \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left\{ \vartheta_t C\left(\boldsymbol{w}_t, \widetilde{\boldsymbol{k}}_t, \boldsymbol{y}_t\right) + \sum_{m=1}^{M} \left[q_{m,t} \widetilde{k}_{m,t} + G_m \left(\widetilde{k}_{m,t} - (1-\delta_m) \widetilde{k}_{m,t-1} \right) \right] \right\}, \tag{2.5}$$

with all endogenous choice variables measuring actual quantities employed in the production and, thus, the optimization objective being reflective of actual costs *inclusive* of inefficiencies.

Optimization with Endogenous Inefficiency. While already an improvement over the more traditional framework, the model in (2.4)–(2.5) continues to assume that inefficiency is effectively random and that the firm has no impact on the extent of variable input over-use in its production. Now, although oftentimes not stated explicitly, technical inefficiency in the stochastic frontier analysis is assumed to be known to a decision maker at the firm. However, it is rarely made clear whether inefficiency can be changed unless the model specification allows for the "environmental" factors determining it. This raises a few questions. If inefficiency is known to the decision maker, why is she not reducing it? Can inefficiency be freely adjusted or is this a costly adjustment process? What is the optimal inefficiency level, if any?

In practice, not only is the past inefficiency a part of the firm's information set based upon which it makes decisions, but the improvements in efficiency are also subject to the firm's control. We explore this possibility below, wherein technical inefficiency is treated as a choice variable subject to adjustment frictions and the firm can choose it optimally when making dynamic production decisions. To endogenize dynamic efficiency, we therefore further generalize intertemporal expected cost minimization objective as follows:

$$\min_{\widetilde{\boldsymbol{k}}_t,\vartheta_t} \ \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left\{ \vartheta_t C\left(\boldsymbol{w}_t, \widetilde{\boldsymbol{k}}_t, \boldsymbol{y}_t\right) + \sum_{m=1}^{M} \left[q_{m,t}\widetilde{k}_{m,t} + G_m\left(\widetilde{k}_{m,t} - (1-\delta_m)\widetilde{k}_{m,t-1}\right) \right] + H\big(\underbrace{\ln\vartheta_t - \ln\vartheta_{t-1}}_{\mathcal{V}_t}\big) \right\}, \tag{2.6}$$

where $H(\cdot)$ is a function that captures adjustment costs in ϑ_t, i.e., the cost of improvements in the $\boldsymbol{x}_t$-oriented dynamic efficiency; and the dynamic inefficiency evolves according to a controlled log-linear law of motion:

$$\ln\vartheta_t = \mathcal{V}_t + \ln\vartheta_{t-1},$$

with $\mathcal{V}_t$ being the improvement[4] in inefficiency ϑ_t; and $C(\cdot)$ is as defined in (2.4). Thus, dynamic inefficiency plays the role of an additional state variable. Here, we effectively conceptualize ϑ_t akin to a dynamic intangible "input" subject to adjustment frictions that result in delayed changes therein. Analogously, $\mathcal{V}_t$ may be thought of as a "net investment" into improved efficiency. We postulate the law of motion for ϑ_t in logs to transform its codomain to have the evolving variable be non-negative just like other dynamic state variables $\{k_{m,t}\}$.

A few more remarks about our conceptualization of inefficiency in the context of the firm's dynamic production decisions. First, just like in the standard stochastic frontier analysis, we do not model inefficiency as a measure relative to other firms' performance but as a deviation from the firm's own "true frontier." Thus, for the decision maker to know her firm's inefficiency, no knowledge about other firms is required. Second, we assume that the decision maker knows the firm's inefficiency, but the latter is unknown to the econometrician. This inefficiency comes from the firm's failure to attain its production frontier. Such failures may be due to bad luck, machine breakdown, mismanagement, human errors, etc. If she does not know it exactly, the decision maker at the firm is reasonably expected at least to have a good estimate of such inefficiency. Third, production decisions and the efficiency are oftentimes influenced by exogenous contextual factors that the firm takes into consideration when optimizing and that are neither inputs nor outputs. We have abstracted away from such a possibility having instead opted for a simpler model in order to not distract the reader from the main focus of our paper which is on the endogenization of inefficiency. Empirically however, such contextual factors can be seamlessly accounted for by conditioning the variable cost function on them thereby effectively treating them as the "$\boldsymbol{y}$" variables that, just like the output quantities, are exogenous and quasi-fixed under the cost-minimization premise.

Also, note that, for computational tractability of our econometric model, in (2.6) we continue to maintain the assumption that distortions in quasi-fixed factors $\{\eta_{m,t}\}$ are exogenous to the firm albeit, in principle, one can endogenize those as well. As it is to become clearer below, we

[4] That is, improvements in *in*efficiency correspond to *negative* values of $\mathcal{V}_t$.

do nonetheless allow $\{\eta_{m,t}\}$ to also be time-persistent thereby indirectly accommodating dynamic properties therein.

The key feature of our model is that it treats technical inefficiency ϑ_t (the adjustments therein, to be precise) as an additional "choice" variable which generates the optimal path of ϑ_t compatible with the firm's intertemporal cost minimization, thereby enhancing its structurally meaningful interpretation. The optimization problem also implies that $\{\eta_{m,t}\}$, which impact the actual use of quasi-fixed inputs $\{\widetilde{k}_{m,t}\}$, cannot be independent over time.

The Euler equations corresponding to the problem in (2.6) are

$$0 = \vartheta_t \frac{\partial C(\cdot)}{\partial \widetilde{k}_{m,t}} + q_{m,t} + \frac{\partial G_m(\cdot)}{\partial i_{m,t}} - \beta(1-\delta_m)\mathbb{E}_t \left\{ \frac{\partial G_m(\cdot)}{\partial i_{m,t+1}} \right\} \qquad \forall\ m = 1, \ldots, M \tag{2.7}$$

$$0 = \vartheta_t C(\cdot) + \frac{\partial H(\cdot)}{\partial \mathcal{V}_t} - \beta \mathbb{E}_t \left\{ \frac{H(\cdot)}{\partial \mathcal{V}_{t+1}} \right\}, \tag{2.8}$$

along with the (actual) conditional demand equations for static inputs from (2.4):

$$x_{j,t} = \vartheta_t \frac{\partial C(\cdot)}{\partial w_{j,t}} \qquad \forall\ j = 1, \ldots, J. \tag{2.9}$$

Note that, in our setup, both the variable-input-oriented technical inefficiency ϑ_t and the distortions in quasi-fixed inputs $\boldsymbol{\eta}_t$ are the short-run measures because they affect the firm's outputs in time t. Their long-run counterparts would be a solution to the system in (2.7)–(2.9) when all other variables are at their long-run stead-steady values.

For the practical implementation, we make the following parametric functional form assumptions. The variable cost function is said to take the translog form, i.e.,

$$\begin{aligned} \ln C(\boldsymbol{w}_t, \widetilde{\boldsymbol{k}}, \boldsymbol{y}_t) = {} & \boldsymbol{\alpha}_o + \boldsymbol{\alpha}_w' \ln \boldsymbol{w}_t + \tfrac{1}{2} \ln \boldsymbol{w}_t' \boldsymbol{\Gamma}_{ww} \ln \boldsymbol{w}_t \; + \\ & \boldsymbol{\alpha}_k' (\ln \boldsymbol{k}_t + \boldsymbol{\xi}_t) + \tfrac{1}{2} (\ln \boldsymbol{k}_t + \boldsymbol{\xi}_t)' \boldsymbol{\Gamma}_{kk} (\ln \boldsymbol{k}_t + \boldsymbol{\xi}_t) + \ln \boldsymbol{w}_t' \boldsymbol{\Gamma}_{wk} (\ln \boldsymbol{k}_t + \boldsymbol{\xi}_t) \; + \\ & \boldsymbol{\alpha}_y' \ln \boldsymbol{y}_t + \tfrac{1}{2} \ln \boldsymbol{y}_t' \boldsymbol{\Gamma}_{yy} \ln \boldsymbol{y}_t + \ln \boldsymbol{y}_t' \boldsymbol{\Gamma}_{yw} \ln \boldsymbol{w}_t + \ln \boldsymbol{y}_t' \boldsymbol{\Gamma}_{yk} (\ln \boldsymbol{k}_t + \boldsymbol{\xi}_t) + \ln \vartheta_t \qquad (2.10) \\ \equiv {} & \mathscr{C}(\ln \boldsymbol{w}_t, \ln \boldsymbol{k}_t + \boldsymbol{\xi}_t, \ln \boldsymbol{y}_t; \boldsymbol{\theta}) + \ln \vartheta_t, \qquad (2.11) \end{aligned}$$

where $\boldsymbol{\alpha}_0, \boldsymbol{\alpha}_w, \boldsymbol{\alpha}_k, \boldsymbol{\alpha}_y$ and $\boldsymbol{\Gamma}_{ww}, \boldsymbol{\Gamma}_{kk}, \boldsymbol{\Gamma}_{wk}, \boldsymbol{\Gamma}_{yy}, \boldsymbol{\Gamma}_{yw}, \boldsymbol{\Gamma}_{yk}$ are respectively vectors and (symmetric) matrices of conformable dimensions containing unknown parameters, which we denote collectively by $\boldsymbol{\theta} \in \Re^P$; and

$$\ln \widetilde{\boldsymbol{k}} = \ln(\boldsymbol{\eta}_t \odot \boldsymbol{k}_t) = \ln \boldsymbol{k}_t + \boldsymbol{\xi}_t,$$

with $\boldsymbol{\xi}_t \equiv \ln \boldsymbol{\eta}_t \in \Re^M$ and $\ln \vartheta_t \in \Re_+$ being, respectively, the unobservable quasi-fixed factor distortion and inefficiency terms (in logs).

We assume convex factor adjustment cost functions, for which we adopt the quadratic form which is a popular specification in both the theoretical and empirical literature (e.g., Gould, 1968;

Pindyck & Rotemberg, 1983; Hall, 2004):

$$G_m(i_{m,t}) = \tfrac{1}{2}\gamma_m i_{m,t}^2 \quad \text{with} \quad \gamma_m > 0, \quad \forall\, m = 1, \dots, M \tag{2.12}$$
$$H(\mathcal{V}_t) = \tfrac{1}{2}\gamma_o \mathcal{V}_t^2 \quad \text{with} \quad \gamma_o > 0 \tag{2.13}$$

and, for convenience, we let $\boldsymbol{\gamma} = [\gamma_1, \dots, \gamma_M, \gamma_o]'$.

With this, the system of dynamic and static optimizing conditions in (2.7)–(2.9) then takes the following form (maintaining the order of appearance of equations):

$$\begin{aligned}
0 = {} & \frac{\vartheta_t}{\eta_{m,t}}\Big[\boldsymbol{\alpha}_k + \boldsymbol{\Gamma}_{kk}\left(\ln \boldsymbol{k}_t + \boldsymbol{\xi}_t\right) + \boldsymbol{\Gamma}_{wk} \ln \boldsymbol{w}_t + \boldsymbol{\Gamma}_{yk} \ln \boldsymbol{y}_t\Big] \frac{\exp\left\{\mathscr{C}(\ln \boldsymbol{w}_t, \ln \boldsymbol{k}_t + \boldsymbol{\xi}_t, \ln \boldsymbol{y}_t; \boldsymbol{\theta})\right\}}{k_{m,t}} + \\
& q_{m,t} + \gamma_m\Big(\eta_{m,t}k_{m,t} - (1-\delta_m)\eta_{m,t-1}k_{m,t-1}\Big) - \\
& \gamma_m \beta (1-\delta_m)\mathbb{E}_t\Big\{\eta_{m,t+1}k_{m,t+1} - (1-\delta_m)\eta_{m,t}k_{m,t}\Big\} \qquad \forall\, m = 1, \dots, M
\end{aligned} \tag{2.14}$$

$$0 = \vartheta_t \exp\left\{\mathscr{C}(\ln \boldsymbol{w}_t, \ln \boldsymbol{k}_t + \boldsymbol{\xi}_t, \ln \boldsymbol{y}_t; \boldsymbol{\theta})\right\} + \gamma_o(\ln \vartheta_t - \ln \vartheta_{t-1}) - \gamma_o \beta \mathbb{E}_t\Big\{\ln \vartheta_{t+1} - \ln \vartheta_t\Big\} \tag{2.15}$$

$$\mathbf{v}_t = S_{j,t} - \vartheta_t\Big[\boldsymbol{\alpha}_w + \boldsymbol{\Gamma}_{ww} \ln \boldsymbol{w}_t + \boldsymbol{\Gamma}_{wk}\left(\ln \boldsymbol{k}_t + \boldsymbol{\xi}_t\right) + \boldsymbol{\Gamma}_{yw} \ln \boldsymbol{y}_t\Big] \qquad \forall\, j = 2, \dots, J, \tag{2.16}$$

where the firm's static condition (2.9) is presented in (2.16) in the share form with one of the share equations omitted; $S_{j,t} \equiv \frac{w_{j,t}x_{j,t}}{\boldsymbol{w}_t'\boldsymbol{x}_t}$ denotes the jth input's variable cost share; and we append these share equations with a $(J-1)\times 1$ vector of two-sided mean-zero stochastic disturbances $\mathbf{v}_t$. During the estimation, we also include the variable cost function (2.11), i.e.,

$$v_{1,t} = \ln(\boldsymbol{w}_t'\boldsymbol{x}_t) - \mathscr{C}(\ln \boldsymbol{w}_t, \ln \boldsymbol{k}_t + \boldsymbol{\xi}_t, \ln \boldsymbol{y}_t; \boldsymbol{\theta}) - \ln \vartheta_t, \tag{2.17}$$

mainly to add information about dynamic technical inefficiency ϑ_t. Notably, ϑ_t directly enters all estimating equations (2.14)–(2.17). Further note that, just like with the derived share equations in (2.16), here we also augment the variable cost equation with a two-sided zero-mean random error term $v_{1,t}$, which may correlate with $\mathbf{v}_t$. We do not impose any additional structure on these random disturbances, including distributional assumptions (except for their mean-independence from the predetermined data) which is a great advantage over previous techniques such as that of Tsionas (2006).

Also, note that the above system of simultaneous equations does not contain endogenous static input quantities. More generally, the structural identification of our estimator is rooted in the economic theory yielding the moment conditions that utilize input prices, predetermined quasi-fixed inputs as well as outputs (under the competitive cost-minimizing paradigm) as sources of weekly exogenous variation. With this, we can compactly rewrite the $(M + J + 1)$ estimating equations in (2.14)–(2.17) as a simultaneous system of moments:

$$\mathbb{E}_t\mathcal{F}(\boldsymbol{\lambda}_{t-1}, \boldsymbol{\lambda}_t, \boldsymbol{\lambda}_{t+1}, \mathbf{Y}_t, \boldsymbol{\Theta}) = \mathbf{0}_{M+J+1}, \tag{2.18}$$

where $\mathcal{F}(\cdot)$ is a vector function of observable data $\mathbf{Y}_t = [\ln \boldsymbol{w}_t', \ln \boldsymbol{y}_t', \boldsymbol{k}_{t-1}', \boldsymbol{k}_t', \boldsymbol{k}_{t+1}', \boldsymbol{q}_t]'$, unknown parameters $\boldsymbol{\Theta} = [\beta, \boldsymbol{\delta}', \boldsymbol{\gamma}', \boldsymbol{\theta}']'$ and latent inefficiency and distortions $\boldsymbol{\lambda}_t = [\vartheta_t, \boldsymbol{\xi}_t']'$, containing the right-hand sides of the Euler equations (2.14)–(2.15), stochastic share equations (2.16) and the stochastic variable cost function (2.17). Note that, if it were not for the unobserved ϑ_t and $\{\xi_{m,t}\}$ inside the moment conditions, equations in (2.18) could have been estimated by the multiple-equation Generalized Method of Moments (GMM). In our case however, things are not as computationally simple (also see Gallant et al., 2017; Gallant, Hong & Khwaja, 2018).

Before proceeding to the discussion of estimation of the system in (2.18), we would like to note that, in our model, the econometric endogenization of inefficiency is firmly aligned with the economic notion of an endogenous variable in the sense that it is a choice variable. That is, the empirical treatment of endogenous ϑ_t does not reduce to merely modeling the atheoretic correlation between this unobservable and regressors. In fact, under our structural assumptions about firm behavior, ϑ_t is conceptualized analogously to a dynamic "input" subject to adjustment frictions which renders it predetermined and thus weakly mean-independent of all regressors in our system of estimating equations in (2.14)–(2.17). In practice, our empirical endogenization of ϑ_t stems from the explicit inclusion of the intertemporal optimality condition for the former [eq. (2.15)] in the system of estimated simultaneous equations. Also, by letting the inefficiency path be optimally chosen within the firm's dynamic optimization problem and explicitly incorporating the latter in the estimation procedure, we are able to obtain inefficiency estimates that can be meaningfully interpreted as being "dynamically optimal."

3 Moment-Based Empirical Likelihood Estimation with Latent Variables

We estimate our model via a Bayesian Exponentially Tilted Empirical Likelihood (BETEL) method proposed by Schennach (2005) as an alternative to fully parametric Bayesian methods which we modify to accommodate the presence of dynamic latent variables—namely, technical efficiency and the distortions in quasi-fixed factors—in the moment conditions.

To fix ideas, first suppose that no latent variables are involved in the model and we have the moment conditions of the following form: $\mathbb{E}_t\mathcal{G}(\mathbf{Y}_t, \boldsymbol{\Theta}) = \mathbf{0}_{\dim(\mathcal{G})}\ \forall\ t = 1, \ldots, n$, where $\mathbf{Y}_t$ and $\boldsymbol{\Theta}$ respectively represent data for each t and unknown parameters. Also, let the entire data for all $t = 1, \ldots, n$ be denoted by $\mathbf{Y}$. The Bayesian posterior corresponding to the BETEL is given by

$$p(\boldsymbol{\Theta}|\mathbf{Y}) \propto p(\boldsymbol{\Theta}) \prod_{t=1}^{n} \omega_t^*(\boldsymbol{\Theta}), \tag{3.1}$$

where $p(\boldsymbol{\Theta})$ is a prior and $\{\omega_t^*(\boldsymbol{\Theta}),\ t = 1, \ldots, n\}$ are solutions to the following problem:

$$\max_{\{\omega_t\}_{t=1}^n} \quad -\sum_{t=1}^{n} \omega_t \log \omega_t \tag{3.2}$$

$$\text{subject to} \quad \sum_{t=1}^{n} \omega_t = 1 \tag{3.3}$$

$$\sum_{t=1}^{n} \omega_t \mathcal{G}(\mathbf{Y}_t, \boldsymbol{\Theta}) = \mathbf{0}_{\dim(\mathcal{G})}, \tag{3.4}$$

provided that the interior of the convex hull of $\bigcup_{t=1}^{n}\{\mathcal{G}(\mathbf{Y}_t, \boldsymbol{\Theta})\}$ contains the origin.

Now suppose that the model contains dynamic latent variables $\boldsymbol{\lambda}_t$ and we have the moment conditions in (2.18): $\mathbb{E}_t\mathcal{F}(\boldsymbol{\lambda}_{t+1}, \mathbf{Y}_t, \boldsymbol{\Theta}) = \mathbf{0}_{M+J+1}\ \forall\ t = 1, \ldots, n$, where we have suppressed the dependence on $\boldsymbol{\lambda}_{t-1}$ and $\boldsymbol{\lambda}_t$ for notational simplicity. Also, assume that dynamic latent variables evolve according to some autoregressive process:

$$\boldsymbol{\lambda}_{t+1} = m(\boldsymbol{\lambda}_t, \boldsymbol{\pi}) + \varepsilon_t, \tag{3.5}$$

where $m(\cdot)$ is the conditional mean function, $\boldsymbol{\pi}$ is a vector of parameters, and ε_t is a random innovation.

Our objective is to reduce our estimation problem, which contains latent variables, into the more conventional BETEL problem so that the posterior result from above may also be used in our case. Thus, our posterior has the following form:

$$p(\boldsymbol{\Theta}, \boldsymbol{\pi}, \boldsymbol{\lambda}|\mathbf{Y}) \propto p(\boldsymbol{\Theta})p(\boldsymbol{\pi}) \prod_{t=1}^{n} p(\boldsymbol{\lambda}_{t+1}|\boldsymbol{\lambda}_t, \boldsymbol{\pi}) \prod_{t=1}^{n} \omega_t^*(\boldsymbol{\Theta}, \boldsymbol{\lambda}), \tag{3.6}$$

where $\boldsymbol{\lambda} = \{\boldsymbol{\lambda}_t, t = 1, ..., n\}$, and $\omega_t^*(\boldsymbol{\Theta}, \boldsymbol{\lambda})$ solves

$$\max_{\{\omega_t\}_{t=1}^{n}} \quad -\sum_{t=1}^{n} \omega_t \log \omega_t \tag{3.7}$$

$$\text{subject to} \quad \sum_{t=1}^{n} \omega_t = 1 \tag{3.8}$$

$$\sum_{t=1}^{n} \omega_t \mathcal{F}(\boldsymbol{\lambda}_{t+1}, \mathbf{Y}_t, \boldsymbol{\Theta}) \otimes \mathbf{z}_t = \mathbf{0}_{(M+J+1)\times\dim(\mathbf{z})}, \tag{3.9}$$

with $\mathbf{z}_t$ being a vector of instruments.

Posterior in (3.6) depends on parameters $\boldsymbol{\Theta}$ and $\boldsymbol{\pi}$ as well as the dynamic latent variables in $\boldsymbol{\lambda}$. While these latent variables are of fundamental interest in themselves, at the same time, they must also be integrated out of the posterior to perform statistical inference on the parameters:

$$p(\boldsymbol{\Theta}, \boldsymbol{\pi}|\mathbf{Y}) \propto \int p(\boldsymbol{\Theta}, \boldsymbol{\pi}, \boldsymbol{\lambda}|\mathbf{Y}) d\boldsymbol{\lambda}, \tag{3.10}$$

but the integral, in general, is impossible to evaluate analytically.

Before proceeding further, we first need to specify the process in (3.5). We opt for a second-order vector autoregressive (VAR) specification with Gaussian innovations, i.e.,

$$\boldsymbol{\Lambda}_t = \boldsymbol{\pi}_0 + \boldsymbol{\pi}_1 \boldsymbol{\Lambda}_{t-1} + \boldsymbol{\pi}_2 \boldsymbol{\Lambda}_{t-2} + \boldsymbol{\varepsilon}_t \quad \text{with} \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}_{M+1}, \boldsymbol{\Sigma}_{\varepsilon}), \tag{3.11}$$

where $\boldsymbol{\Lambda}_t \equiv [\ln \frac{1}{\vartheta_t - 1}, \boldsymbol{\xi}_t']'$ contains (a function of) dynamic technical inefficiency $\vartheta_t \geq 1$ and the distortion factors pertaining to the quasi-fixed inputs $\{\xi_{m,t}\}$; $\boldsymbol{\pi}_0$, $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$ are the $(M+1) \times 1$ parameter vectors and $\boldsymbol{\pi} = [\boldsymbol{\pi}_0', \boldsymbol{\pi}_1', \boldsymbol{\pi}_2']'$. The transformation of ϑ_t into $\ln \frac{1}{\vartheta_t - 1}$ simplifies estimation by having the latter be defined on the whole real line. The choice of a second-order VAR model is motivated by the second-order dynamics induced by the optimization problem; see eq. (2.15). Initial conditions for the latent variables, i.e., ϑ_0 and $\boldsymbol{\xi}_0$, are treated as being unknown to the firm (as parameters).

We use Markov Chain Monte Carlo (MCMC) methods to perform computations. Our MCMC involves two steps that are carried out for each MCMC iteration. In the first step, we use Sequential Monte Carlo (SMC), or Particle Filtering (PF), to provide draws for $\{\boldsymbol{\lambda}_t^{(i)}, i = 1, \ldots, N\}$, where i indexes the MCMC simulation, and N is the total number of such simulations. In the second step, we draw parameters $\boldsymbol{\Theta}^{(i)}$ and $\boldsymbol{\pi}^{(i)}$. Since standard Metropolis-Hastings algorithms may be quite computationally inefficient, we use Girolami & Calderhead's (2011) Riemann manifold Langevin–Hamiltonian Monte Carlo method (hereafter, the GC algorithm). This technique is reliable, requires almost no tuning, and the MCMC draws that it provides have considerably less autocorrelation compared to other MCMC algorithms.

In what follows, we briefly describe the employed methodologies.

Step 1. The SMC/PF methodology is applied to state-space models of the following generic form:

$$\mathbf{Y}_{1:T} \sim p(\mathbf{Y}_t | \boldsymbol{\lambda}_t) \quad \text{and} \quad \boldsymbol{\lambda}_t \sim p(\boldsymbol{\lambda}_t | \boldsymbol{\lambda}_{t-1}), \tag{3.12}$$

where $\boldsymbol{\lambda}_t$ are state variables (for the background on Bayesian state estimation, see Gordon, Salmond & Smith, 1993; Gordon, 1997; Doucet, Freitas & Gordon, 2001; Ristic, Gordon & Arulampalam, 2004). Given the data $\mathbf{Y}_t$, the posterior distribution $p(\boldsymbol{\lambda}_t | \mathbf{Y}_t)$ can be approximated by a set of (auxiliary) particles $\{\boldsymbol{\lambda}_t^{(i)}, i = 1, \ldots, N\}$ with probability weights $\{w_t^{(i)}, i = 1, \ldots, N\}$ such that $\sum_{i=1}^{N} w_t^{(i)} = 1$. With this, we can approximate the predictive density by

$$p(\boldsymbol{\lambda}_{t+1} | \mathbf{Y}_t) = \int p(\boldsymbol{\lambda}_{t+1} | \boldsymbol{\lambda}_t) p(\boldsymbol{\lambda}_t | \mathbf{Y}_t) d\boldsymbol{\lambda}_t \simeq \sum_{i=1}^{N} p(\boldsymbol{\lambda}_{t+1} | \boldsymbol{\lambda}_t^{(i)}) w_t^{(i)}, \tag{3.13}$$

with the final approximation for the filtering density being given by

$$p(\boldsymbol{\lambda}_{t+1} | \mathbf{Y}_t) \propto p(\mathbf{Y}_{t+1} | \boldsymbol{\lambda}_{t+1}) p(\boldsymbol{\lambda}_{t+1} | \mathbf{Y}_t) \simeq p(\mathbf{Y}_{t+1} | \boldsymbol{\lambda}_{t+1}) \sum_{i=1}^{N} p(\boldsymbol{\lambda}_{t+1} | \boldsymbol{\lambda}_t^{(i)}) w_t^{(i)}. \tag{3.14}$$

Then, the basic mechanism of particle filtering rests on propagating $\{\boldsymbol{\lambda}_t^{(i)}, w_t^{(i)}, i = 1, \ldots, N\}$ to the next step, i.e., $\{\boldsymbol{\lambda}_{t+1}^{(i)}, w_{t+1}^{(i)}, i = 1, \ldots, N\}$. Given $\boldsymbol{\Theta}$, Andrieu & Roberts (2009), Flury & Shephard (2011) and Pitt, dos Santos Silva, Giordani & Kohn (2012) provide the Particle Metropolis-Hastings (PMCMC) technique which uses an unbiased estimator of the likelihood function $\widehat{p}_N(\mathbf{Y}|\boldsymbol{\Theta})$ since $p(\mathbf{Y}|\boldsymbol{\Theta})$ is often unavailable in a closed form.

We use Gordon et al.'s (1993) particle filter, where particles are simulated through the state density $p(\boldsymbol{\lambda}_t^{(i)}|\boldsymbol{\lambda}_{t-1}^{(i)})$ and then re-sampled with weights determined by the measurement density evaluated at the resulting particle, i.e., $p(\mathbf{Y}_t|\boldsymbol{\lambda}_t^{(i)})$. The latter is simple to construct and rests upon the following steps, for $t = 0, \ldots, T-1$ given samples $\boldsymbol{\lambda}_t^k \sim p(\boldsymbol{\lambda}_t|\mathbf{Y}_{1:t})$ with mass π_t^k for $k = 1, \ldots, N$:

1. For $k = 1, \ldots, N$, compute $\omega_{t|t+1}^{(k)} = g\left(\mathbf{Y}_{t+1}|\boldsymbol{\lambda}_t^{(k)}\right)\pi_t^{(k)}$, $\pi_{t|t+1}^{(k)} = \omega_{t|t+1}^{(k)} \Big/ \sum_{i=1}^N \omega_{t|t+1}^{(i)}$.
2. For $k = 1, \ldots, N$, draw $\widetilde{\boldsymbol{\lambda}}_t^{(k)} \sim \sum_{i=1}^N \pi_{t|t+1}^{(i)} \delta_{\boldsymbol{\lambda}_t}^{(i)}(d\boldsymbol{\lambda}_t)$.
3. For $k = 1, \ldots, N$, draw $u_{t+1}^{(k)} \sim g\left(u_{t+1}\middle|\widetilde{\boldsymbol{\lambda}}_t^{(k)}, \mathbf{Y}_{t+1}\right)$ and set $\boldsymbol{\lambda}_{t+1}^{(k)} = h\left(\boldsymbol{\lambda}_t^{(k)}; u_{t+1}^{(k)}\right)$.
4. For $k = 1, \ldots, N$, compute

$$\omega_{t+1}^{(k)} = \frac{p\left(\mathbf{Y}_{t+1}|\boldsymbol{\lambda}_{t+1}^{(k)}\right) p\left(u_{t+1}^{(k)}\right)}{g\left(\mathbf{Y}_{t+1}|\boldsymbol{\lambda}_t^{(k)}\right) g\left(u_{t+1}^{(k)}|\widetilde{\boldsymbol{\lambda}}_t^{(k)}, \mathbf{Y}_{t+1}\right)} \quad \text{and} \quad \pi_{t+1}^{(k)} = \frac{\omega_{t+1}^{(k)}}{\sum_{i=1}^N \omega_{t+1}^{(i)}}.$$

Lastly, the estimate of likelihood from ADPF is $p(\mathbf{Y}_{1:T}) = \prod_{t=1}^T \left(\sum_{i=1}^N \omega_{t-1|t}^{(i)}\right)\left(N^{-1}\sum_{i=1}^N \omega_t^{(i)}\right)$.

Step 2. To update draws for the parameter vector of interest $\boldsymbol{\delta} = (\boldsymbol{\Theta}', \boldsymbol{\pi}')'$, the GC algorithm uses local information about both the gradient and the Hessian of the log-posterior conditional on $\boldsymbol{\delta}$ at the existing draw. The GC algorithm is started at the first-stage GMM estimator and MCMC is run until convergence. Depending on the model and the subsample, this takes 5,000 to 10,000 iterations. We opt for 10,000 iterations. Then, we run another 50,000 MCMC iterations to obtain the final results for posterior moments and densities of parameters and functions of interest.

Let $\mathscr{L}(\boldsymbol{\delta}) = \log p(\boldsymbol{\delta}|\mathbf{Y})$ denote the log posterior of $\boldsymbol{\delta}$. Also, define $V(\boldsymbol{\delta}) = \text{est.cov}\frac{\partial}{\partial \boldsymbol{\delta}} \log p(\mathbf{Y}|\boldsymbol{\delta})$ to be the empirical counterpart of $V_o(\boldsymbol{\delta}) = -\mathbb{E}_{\mathbf{Y}|\boldsymbol{\delta}} \frac{\partial^2}{\partial \boldsymbol{\delta} \partial \boldsymbol{\delta}'} \log p(\mathbf{Y}|\boldsymbol{\delta})$. The Langevin diffusion for the parameter vector $\boldsymbol{\delta}$ is given by the following stochastic differential equation:

$$d\boldsymbol{\delta}(t) = \tfrac{1}{2}\widetilde{\nabla}_{\boldsymbol{\delta}}\mathscr{L}\{\boldsymbol{\delta}(t)\}\,dt + dB(t), \tag{3.15}$$

where

$$\widetilde{\nabla}_{\boldsymbol{\delta}}\mathscr{L}\{\boldsymbol{\delta}(t)\} = -V^{-1}\{\boldsymbol{\delta}(t)\} \cdot \nabla_{\boldsymbol{\delta}}\mathscr{L}\{\boldsymbol{\delta}(t)\} \tag{3.16}$$

is the "natural gradient" of the Riemann manifold corresponding to the log posterior. The elements of the Brownian motion in (3.15) are

$$V^{-1}\left\{\boldsymbol{\delta}\left(t\right)\right\}dB_i\left(t\right)=\left|V\left\{\boldsymbol{\delta}\left(t\right)\right\}\right|^{-1/2}\sum_{j=1}^{\dim(\boldsymbol{\delta})}\tfrac{\partial}{\partial\boldsymbol{\delta}}\left[V^{-1}\left\{\boldsymbol{\delta}\left(t\right)\right\}_{ij}\left|V\left\{\boldsymbol{\delta}\left(t\right)\right\}\right|^{1/2}\right]dt+\left\{\sqrt{V\left\{\boldsymbol{\delta}\left(t\right)\right\}}dB\left(t\right)\right\}_i. \tag{3.17}$$

The discrete form of the stochastic differential equation in (3.17) provides a proposal as follows:

$$\begin{aligned}\widetilde{\boldsymbol{\delta}}_i=\boldsymbol{\delta}_i^o&+\tfrac{\varepsilon^2}{2}\left\{V^{-1}\left(\boldsymbol{\delta}^o\right)\nabla_{\boldsymbol{\delta}}\mathscr{L}\left(\boldsymbol{\delta}^o\right)\right\}_i-\varepsilon^2\sum_{j=1}^{\dim(\boldsymbol{\delta})}\left\{V^{-1}\left(\boldsymbol{\delta}^o\right)\frac{\partial V\left(\boldsymbol{\delta}^o\right)}{\partial\boldsymbol{\delta}j}V^{-1}\left(\boldsymbol{\delta}^o\right)\right\}_{ij}+\\&\tfrac{\varepsilon^2}{2}\sum_{j=1}^{\dim(\boldsymbol{\delta})}\left\{V^{-1}\left(\boldsymbol{\delta}^o\right)\right\}_{ij}\operatorname{tr}\left\{V^{-1}\left(\boldsymbol{\delta}^o\right)\frac{\partial V\left(\boldsymbol{\delta}^o\right)}{\partial\boldsymbol{\delta}j}\right\}+\left\{\varepsilon\sqrt{V^{-1}\left(\boldsymbol{\delta}^o\right)}\xi^o\right\}_i\\&\equiv\mu\left(\boldsymbol{\delta}^o,\varepsilon\right)_i+\left\{\varepsilon\sqrt{V^{-1}\left(\boldsymbol{\delta}^o\right)}\xi^o\right\}_i,\end{aligned} \tag{3.18}$$

where $\boldsymbol{\delta}^o$ is the current draw, and ϵ is selected so that accept rate is about 25%. The proposal density is

$$q\left(\widetilde{\boldsymbol{\delta}}\middle|\boldsymbol{\delta}^o\right)=\mathcal{N}_{\dim(\boldsymbol{\delta})}\left(\widetilde{\boldsymbol{\delta}},\varepsilon^2V^{-1}\left(\boldsymbol{\delta}^o\right)\right), \tag{3.19}$$

and convergence to the invariant distribution is ensured by using the standard-form Metropolis-Hastings probability:

$$\min\left\{1,\ \frac{p\left(\widetilde{\boldsymbol{\delta}}\middle|\cdot,\mathbf{Y}\right)q\left(\boldsymbol{\delta}^o\middle|\widetilde{\boldsymbol{\delta}}\right)}{p\left(\boldsymbol{\delta}^o\middle|\cdot,\mathbf{Y}\right)q\left(\widetilde{\boldsymbol{\delta}}\middle|\boldsymbol{\delta}^o\right)}\right\}. \tag{3.20}$$

4 Empirical Application

Data. We use annual state-level panel data on agricultural farm production in lower 48 contiguous states of the United States during the 1960–2004 period reported by the Economic Research Service (ERS) of the U.S. Department of Agriculture. These data, or subsets thereof, have been used in the literature before, e.g., by O'Donnell (2012, 2016). There are three outputs—crops (y_1), livestock (y_2) and other outputs (y_3)—and four inputs: capital (k_1), land (k_2), labor (x_1) and total intermediate inputs (x_2). The ERS constructs these output/input accounts for the farm sector consistent with a gross output model of production [for details, see Ball, Bureau, Nehring & Somwaru (1997); Ball, Gollop, Kelly-Hawke & Swinand (1999), or the ERS methods manual]. There is also information on prices. All data are in the form of relative indices. Physical capital is subject to adjustment frictions (e.g., time-to-install) and thus is quasi-fixed. It is perhaps even more natural to assume that land is quasi-fixed too. Labor and intermediate inputs are however treated as freely varying. For capital and land, we allow for unknown depreciation rates.[5] The

[5]As pointed out by a referee, assuming depreciation of land may be rather counter-intuitive because land has no definitive useful life. We opt for a non-zero depreciation rate here primarily to allow for a greater degree of flexibility

time discount factor β is also treated as an unknown parameter.

Estimation Details. Our priors are as follows. For the discount factor β, we use a beta prior $\mathcal{B}(\underline{a}, \underline{b})$. The same prior is used for the input adjustment cost coefficients $\{\gamma_m,\ m = 1, \ldots, 2\}$, the efficiency adjustment parameter γ_o, and the depreciation rate parameters for capital and land $\{\delta_m,\ m = 1, 2\}$. In our baseline specification, we set $\underline{a} = 5$ and $\underline{b} = 1$. For the translog parameters in (2.11), we adopt a flat prior subject to the restriction that monotonicity conditions hold at the sample means and ten other randomly selected points in the hypercube defined by the posterior means plus/minus twice the sample standard deviations. We do not impose monotonicity conditions at other sample points but examine whether they hold at them (using the posterior means of the parameters) once MCMC is finished. For the $\boldsymbol{\pi}$ parameters of the VAR model in (3.11), our prior is $\mathcal{N}_3(\underline{\boldsymbol{\pi}}, \underline{h}\mathbf{I}_3)$, where $\underline{\boldsymbol{\pi}} = \mathbf{0}_3$ and $\underline{h} = 1$. Using a relatively small value for $\underline{h}$ implies considerable shrinkage (although this can be detected only if we know the likelihood contribution). For the covariance matrix of the VAR errors $\boldsymbol{\Sigma}_\varepsilon$, our prior is $p(\boldsymbol{\Sigma}_\varepsilon) \propto |\boldsymbol{\Sigma}_\varepsilon|^{-(\underline{n}+1)} \exp\{-\underline{\mathbf{A}}\boldsymbol{\Sigma}_\varepsilon^{-1}\}$, where we set $\underline{n} = 1$ and $\underline{\mathbf{A}} = \underline{a}_2\mathbf{I}_3$ with $\underline{a}_2 = 0.1$. For the initial conditions $\boldsymbol{\Lambda}_1$ and $\boldsymbol{\Lambda}_2$ in (3.11), our prior is $\mathcal{N}_{M+1}(\mathbf{0}_3, \underline{h}_2\mathbf{I}_3)$ with $\underline{h}_2 = 10$. Alternative prior specifications are considered by varying hyper-parameters $\underline{a}$, $\underline{b}$, $\underline{\boldsymbol{\pi}}$, $\underline{h}$, $\underline{n}$, $\underline{a}_2$ and $\underline{h}_2$. For sensitivity analysis, all these parameters are multiplied by uniform random numbers from the $(0, 100)$ interval. In the case of $\underline{\boldsymbol{\pi}}$, the resulting hyper-parameters are also assigned a negative sign with the probability 1/2.

In the estimation, we impose the symmetry and linear homogeneity (in input prices) restrictions onto the dual cost function in (2.17) which, naturally, also imply restrictions for the remaining equations (2.14)–(2.16) in the system. Further, to allow for temporal shifts in the technological frontier, we also include a time trend (along with its square and interactions terms) in the translog cost function $\mathscr{C}(\cdot)$ as well as a series of state dummy variables capturing unobservable state fixed effects. Drawing on a structural identification power of the cost-minimization premise in face of perfectly competitive factor markets, we use logs of input prices, quasi-fixed input and output quantities along with the time trend and state dummies for instruments $\mathbf{z}_t$ in (3.9). Except for dummy variables, we also employ squares and interactions of all these instruments.

Lastly, we implement BETEL using 150,000 MCMC iterations the first 50,000 of which are discarded to mitigate possible start-up effects. The SMC is implemented using 10^{12} particles per MCMC iteration.

Results. To highlight the merits of our proposed methodology as well as to assess sensitivity of the empirical results (and the conclusions that researchers may draw upon them) to the fashion in which dynamic efficiency is conceptualized and modeled—as the firm's endogenous choice variable or as following an exogenous process—we estimate two models:

in our model and to accommodate the possibility of land quality deterioration due to intensive use of farmland or over-use of chemical fertilizers, etc. Having said that, our estimate of the annual depreciation rate for land is indeed very near zero (0.0015; see Table A.1) implying that fixing the former at zero *a priori* would have changed the results little.

Table 1. Posterior estimates of technological metrics

		Model I		Model II	
		Mean	*S.d.*	*Mean*	*S.d.*
Panel A: (Short-Run) Cost Function Elasticities					
Labor Price Elasticity	$\partial\mathcal{C}/\partial \ln w_1$	0.712	(0.022)	0.688	(0.015)
Intermediates Price Elasticity	$\partial\mathcal{C}/\partial \ln w_2$	0.288	(0.022)	0.312	(0.015)
Capital Elasticity	$\partial\mathcal{C}/\partial \ln k_1$	0.227	(0.015)	0.128	(0.027)
Land Elasticity	$\partial\mathcal{C}/\partial \ln k_2$	0.533	(0.031)	0.142	(0.016)
Crop Output Elasticity	$\partial\mathcal{C}/\partial \ln y_1$	0.505	(0.021)	0.368	(0.019)
Livestock Output Elasticity	$\partial\mathcal{C}/\partial \ln y_2$	0.191	(0.016)	0.372	(0.015)
Other Output Elasticity	$\partial\mathcal{C}/\partial \ln y_3$	0.233	(0.045)	0.303	(0.030)
Scale Elasticity of Cost	$\sum_q \partial\mathcal{C}/\partial \ln y_q$	0.975	(0.022)	0.934	(0.030)
Panel B: Dynamic Efficiency & Factor Distortions					
Dynamic technical efficiency	$1/\vartheta$	0.933	(0.012)	0.838	(0.033)
Distortion in capital	η_1	1.025	(0.006)	1.025	(0.007)
Distortion in land	η_2	1.048	(0.012)	1.039	(0.012)
Panel C: Productivity Growth Decomposition					
Efficiency change (EC)	$-d\ln\vartheta/dt$	0.0213	(0.0185)	0.0056	(0.0026)
Technical change (TC)	$-\partial\mathcal{C}/\partial t$	0.0261	(0.0072)	0.0057	(0.0110)
Productivity growth	EC+TC	0.0474	(0.0197)	0.0114	(0.0111)

NOTE: Model I endogenizes dynamic inefficiency, whereas Model II treats it as being exogenous.

I. The model with endogenously determined (optimal) dynamic efficiency implied by (2.6) and estimated using a full system of the variable cost function along with dynamic and static optimizing conditions given in (2.14)–(2.17). This is our proposed (and preferred) model.

II. The model with exogenous dynamic inefficiency implied by (2.5) and estimated using the same system of equations *less* the first-order Euler equation corresponding to efficiency improvements in (2.15). This model is in line with a more traditional formulation of the dynamic production subject to inefficiency although, to ensure maximal comparability with Model I, here we continue to maintain a system approach as well as let technical inefficiency explicitly enter the firm's intertemporal optimization conditions as a state variable.

Table 1 summarizes the results from these two models. The reported are posterior means and standard deviations of key technological metrics. For the estimates of underlying structural parameters entering the firm's expected dynamic optimization problem, see the Appendix.

Before proceeding to the discussion of main results concerning dynamic (in)efficiency and factor distortions, we first examine the scale elasticity estimates $\sum_q \partial\mathcal{C}/\partial \ln y_q$. These estimates are of interest because they gauge returns to scale in the agricultural production. Specifically, the production technology is said to exhibit increasing/constant/decreasing returns to scale if the scale elasticity (of cost) is less than/equal to/greater than one. While both models produce a less-than-one posterior mean point estimate of the scale cost elasticity, in the case of Model I the

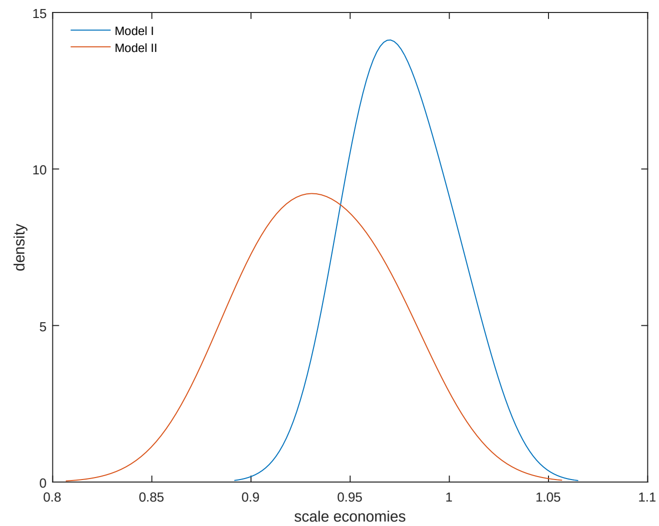


Figure 1. Distributions of posterior estimates of the scale elasticity of cost

corresponding 95% posterior coverage regions includes unity suggesting that, on average, the agricultural farm production sector in the U.S. exhibits constant returns to scale during our sample period, which is consistent with the findings by O'Donnell (2016) who uses similar data for the Northeast in the 1960–1989 period. The alternative Model II however produces qualitatively different evidence in favor of significant economies of scale at the aggregate level. This tendency of the second model to under-estimate scale elasticity and thus to over-estimate returns to scale in the sector is distribution-wise as can vividly be seen in Figure 1, which plots sampling distributions of posterior estimates of the scale elasticity of cost from the two models. Further, the empirical results from the two models also non-negligibly differ in other aspects of the cost relationship. Contrasting the posterior estimates of individual mean cost elasticities (see Panel A of Table 1), we find that treating dynamic efficiency as exogenous (in Model II) appears to dramatically under-estimate sensitivity of costs to quasi-fixed factors as well as to indicate a higher relative importance of intermediate inputs (over labor) in the cost.

Panel B of Table 1 presents the estimates of primary interest to our paper. First, consider the dynamic variable-input-oriented technical efficiency $\vartheta_t^{-1} \in (0, 1]$. We find that the failure to allow for potential endogenous adjustments in efficiency over time (Model II) produces significantly lower estimates of technical efficiency: the pooled mean posterior estimate of 0.84 vs. 0.93 from our endogenous-efficiency Model I. The posterior results from Model II are also significantly more variable, as can be seen from Figure 2, with more than a few point estimates of around 0.8 and lower. Overall, this tendency of Model II to under-estimate efficiency is likely due to its inherent inability to properly credit producing units for incurring efficiency-improvement adjustment costs.[6]

[6]Unfortunately, we cannot analytically derive this "over-estimation of inefficiency" result, which is why we cannot claim the generality of this finding for other applications. Intuitively however, one would expect to obtain larger inefficiency values when using the more standard model of dynamic efficiency that does not endogenize efficiency because, due to its inability to accommodate costly (but optimal) adjustments in efficiency levels that eventually lead to lower observable costs, such a model would not properly credit firms for incurring efficiency-improvement

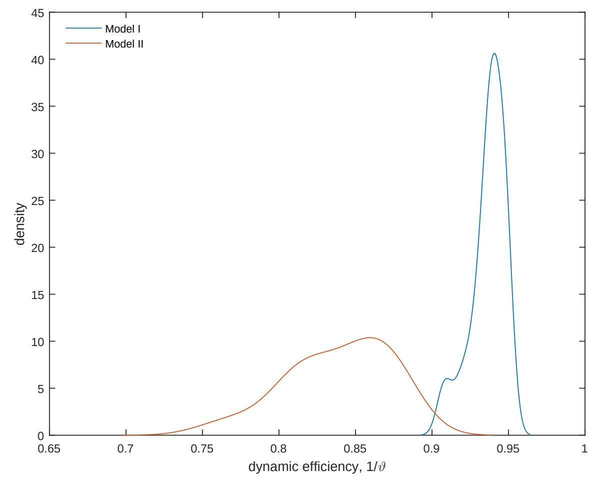
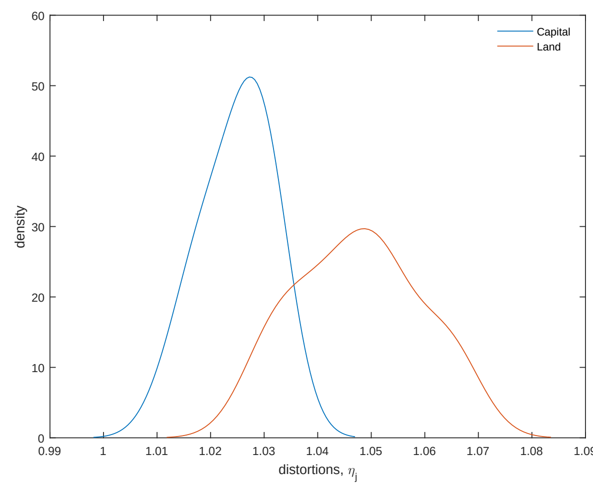
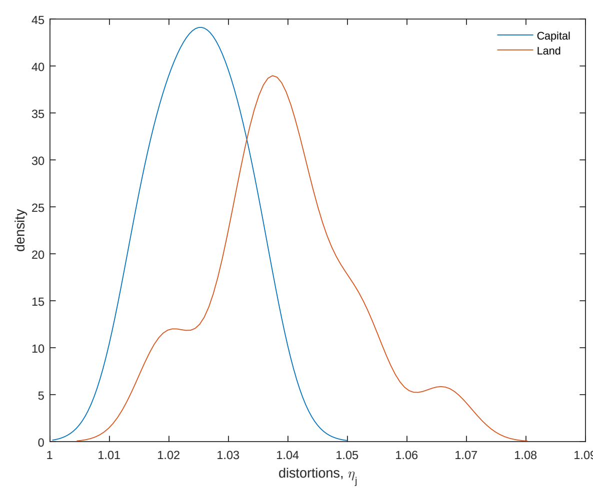


Figure 2. Distributions of posterior estimates of the dynamic efficiency

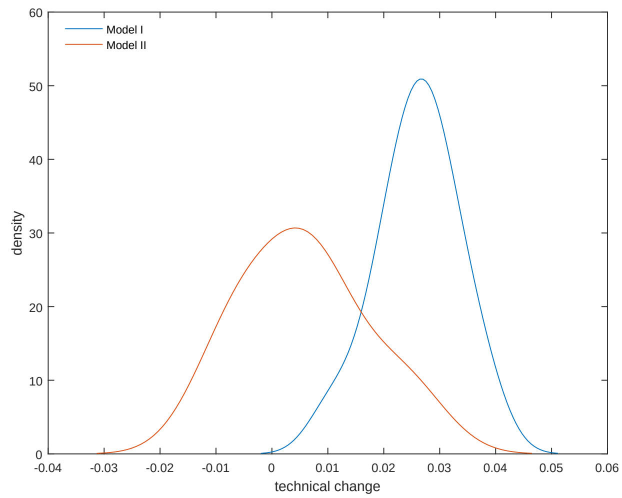


(a) Model I

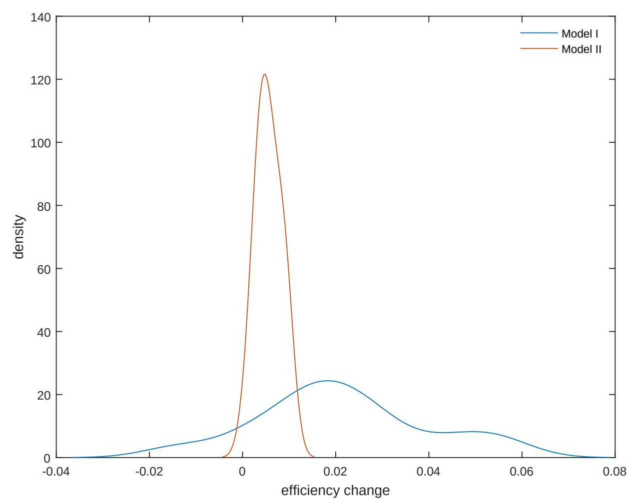


(b) Model II

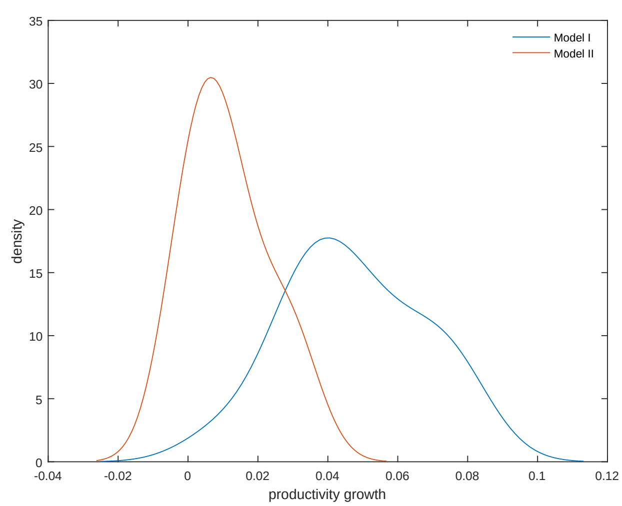
Figure 3. Distributions of posterior estimates of the quasi-fixed factor distortions



(a) Technical change



(b) Efficiency change



(c) Productivity change

Figure 4. Distributions of posterior estimates of the productivity change and its components

Next, we examine distortions in the use of quasi-fixed inputs. Our empirical results suggest several findings. First, regardless of whether the variable-input efficiency treated as exogenous or endogenous, the data consistently indicate an over-use in both quasi-fixed dynamic inputs as evidenced by universally greater-than-one values of posterior estimates of $\eta_1 > 0$ for capital and $\eta_2 > 0$ for land: see Figure 3 that plots their sampling distributions across the two models. Second, both models indicate a greater degree of over-use in land than physical capital. Third, our preferred Model I produces evidence of somewhat greater distortions in the land use, with the posterior mean estimate of 4.8% compared to that of 3.9% from Model II (see Panel B of Table 1). When it comes to the capital use however, the average distortions are the same across the two models (although the distributions thereof are not).

Until now, we have paid little attention to the dynamics in production technology over time. We now specifically focus on temporal changes in agricultural production. The dual cost-function-based productivity growth measure and its decomposition into key components provide a natural avenue for summarizing temporal dynamics in the production. Given our (based on preferred Model I) and the previously reported evidence of unitary returns to scale in the U.S. farm production as well as maintaining our implicit assumption of no allocative inefficiency, the dual multi-output productivity growth index can be easily shown to be a sum of the technical change and efficiency change, with each component respectively defined as $-\partial \ln C_t/\partial t$ and $-d \ln \vartheta_t/dt$ (e.g., see Kumbhakar & Lovell, 2000), which is how we measure it. The posterior mean estimates of productivity growth and its components are reported in Panel C of Table 1; their respective sampling distributions are plotted in Figure 4.

We document stark differences in the productivity growth estimates across the two models. Our preferred endogenous-efficiency Model I estimates the average productivity growth rate in the agricultural farm production at significant 4.7% p.a. whereas the corresponding estimate from the exogenous-efficiency Model II is merely 1.1% and is insignificant. Based on the decomposition results, we find that the latter model produces much smaller close-to-zero estimates of both the mean efficiency and technical change. When endogenizing efficiency however, our results suggest material improvements in dynamic efficiency over time at an about 2.6% average annual rate. Interestingly, our posterior mean estimates of productivity growth in the farm production from Model I are notably greater than those reported earlier by Ball et al. (1997) and more recently by Andersen, Alston, Pardey & Smith (2018) and Plastina & Lence (2018). This may be plausibly attributed to distinct differences in the productivity measurement methodologies. Unlike the cited studies that pursue primal static-production approaches, our methodology is dual and follows a dynamic framework.[7]

Having contrasted empirical results from the two models, a natural question is which of the two is more favored by the data. The posterior estimate of the efficiency adjustment cost parameter

costs.

[7] A more close examination of how the choice of dynamic over static production frameworks affects the measurements of agricultural productivity is beyond the scope of our empirical application but might provide a fruitful avenue for future research.

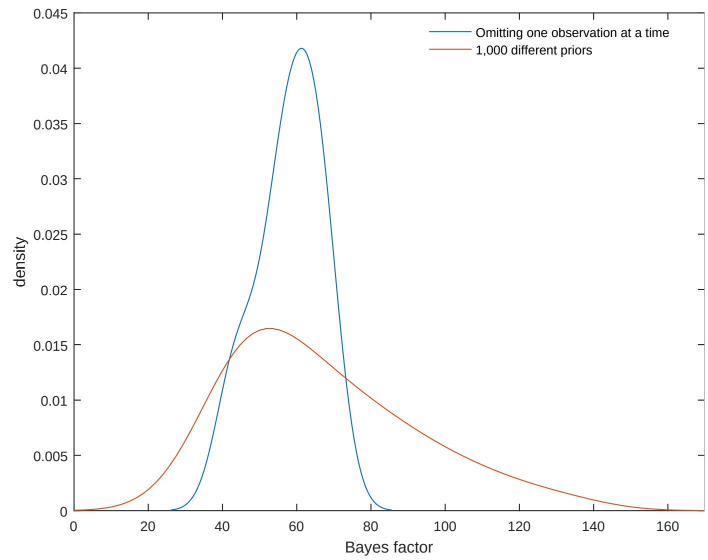


Figure 5. Distributions of the Bayes factor in favor of the endogenous-efficiency model

γ_o from Model I provides an indirect evidence in favor of the endogenous-efficiency model. We estimate it to be significantly above zero (see Table A.1 in the Appendix) which lends support to the dynamic endogenization of technical inefficiency ϑ_t. To select between the two models, we also employ the Bayes factor constructed using their respective posteriors with that of Model II used in the denominator, implying that the values above one would indicate that Model I is more strongly supported by the data. To ensure robustness of model selection to the choice of priors and the outlier influences in the data, we consider 1,000 alternative prior specifications (via changing hyper-parameters) as well as re-estimate the models using leave-one-out subsamples. Figure 5 plots the corresponding sampling distributions of the Bayes factor, from where it is evident that data overwhelmingly favor our preferred specification that endogenizes dynamic efficiency.

5 Conclusion

Existing methods for the measurement of technical efficiency in the dynamic production models obtain it from implied distance functions without making use of the information about intertemporal economic behavior in the estimation beyond an indirect appeal to duality despite seeking a deeply structural "dynamic" interpretation of efficiency. The main limitation of such an estimation approach is that it does not allow for the firm-controlled dynamic evolution of efficiency, thereby effectively assuming that the firm's efficiency is exogenous. In this paper, we introduce a new (structural) conceptualization of efficiency that directly enters the firm's intertemporal production decisions and is both explicitly costly and endogenously determined. The obtained measure of firm efficiency is thus explicitly dynamic. We differentiate between the variable-inputs-oriented inefficiency and factor-specific distortions in quasi-fixed inputs and allow all of these dynamic latent variables to jointly evolve over time. We build a moment-based multiple-equation system estimation procedure that incorporates the variable cost function and both the dynamic and static optimality conditions derived from the firm's intertemporal expected cost minimization. We operationalize our methodology using a modified version of a nonparametric Bayesian Exponentially Tilted Empirical Likelihood adjusted for the presence of dynamic latent variables in the model, which we showcase using the 1960–2004 U.S. agricultural farm production data. Among other things, we find that the failure to allow for potential endogenous adjustments in efficiency over time produces significantly lower estimates of dynamic efficiency, which is likely due to the inherent inability of a more traditional exogenous-efficiency framework to properly credit producing units for incurring efficiency-improvement adjustment costs.

Appendix

Table A.1. Posterior Estimates of Deep Structural Parameters

		Model I		**Model II**	
		Mean	*S.d.*	*Mean*	*S.d.*
Time discount parameter	β	0.710	(0.012)	0.887	(0.015)
Depreciation rate for capital	δ_1	0.044	(0.012)	0.030	(0.007)
Depreciation rate for land	δ_2	0.0015	(0.0004)	0.017	(0.007)
Adjustment cost parameter for capital	γ_1	0.133	(0.021)	0.244	(0.015)
Adjustment cost parameter for land	γ_2	0.732	(0.025)	1.232	(0.120)
Adjustment cost parameter for inefficiency	γ_o	0.255	(0.036)	—	—

NOTE: Model I endogenizes dynamic inefficiency, whereas Model II treats it as being exogenous.

References

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

Andersen, M. A., Alston, J. M., Pardey, P. G., & Smith, A. (2018). A century of U.S. farm productivity growth: A surge then a slowdown. *American Journal of Agricultural Economics*, *100*, 1072–1090.

Andrieu, C. & Roberts, G. O. (2009). The pseudo-marginal approach for efficient Monte Carlo computations. *Annals of Statistics*, *37*, 697–725.

Ball, V. E., Bureau, J., Nehring, R., & Somwaru, A. (1997). Agricultural productivity revisited. *American Journal of Agricultural Economics*, *79*, 1045–1063.

Ball, V. E., Gollop, F. M., Kelly-Hawke, A., & Swinand, G. (1999). Patterns of state productivity growth in the U.S. farm sector: Linking state and aggregate models. *American Journal of Agricultural Economics*, *81*, 164–179.

Chen, C.-M. & van Dalen, J. (2010). Measuring dynamic efficiency: Theories and an integrated methodology. *European Journal of Operational Research*, *203*, 749–760.

Cherchye, L., De Rock, B., & Kerstens, P. J. (2018). Production with storable and durable inputs: Nonparametric analysis o intertemporal efficiency. *European Journal of Operational Research*, *270*, 498–513.

Doucet, A., Freitas, N., & Gordon, N. (2001). *Sequential Monte Carlo Methods in Practice*. New York: Springer-Verlag.

Flury, T. & Shephard, N. (2011). Bayesian inference based only on simulated likelihood: Particle filter analysis of dynamic economic models. *Econometric Theory*, *27*, 933–956.

Gallant, A. R., Giocomini, R., & Ragusa, G. (2017). Bayesian estimation of state space models using moment conditions. *Journal of Econometrics*, *201*, 198–211.

Gallant, A. R., Hong, H., & Khwaja, A. (2018). A Bayesian approach to estimation of dynamic models with small and large number of heterogeneous players and latent serially correlated states. *Journal of Econometrics*, *203*, 19–32.

Girolami, M. & Calderhead, B. (2011). Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *Journal of the Royal Statistical Society B*, *73*, 123–214.

Gordon, N. J. (1997). A hybrid bootstrap filter for target tracking in clutter. *IEEE Transactions on Aerospace and Electronic Systems*, *33*, 353–358.

Gordon, N. J., Salmond, D. J., & Smith, A. F. M. (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEEE-Proceedings-F*, *140*, 107–113.

Gould, J. P. (1968). Adjustment costs in the theory of investment of the firm. *Review of Economic Studies*, *35*, 47–55.

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

Hampf, B. (2017). Rational inefficiency, adjustment costs and sequential technologies. *European Journal of Operational Research*, *263*, 1095–1108.

Kapelko, M. & Oude Lansink, A. (2017). Dynamic multi-directional inefficiency analysis of European diary manufacturing firms. *European Journal of Operational Research*, *257*, 338–344.

Kapelko, M., Oude Lansink, A., & Stefanou, S. E. (2014). Assessing dynamic inefficiency of the Spanish construction sector pre- and post-financial crisis. *European Journal of Operational Research*, *237*, 349–357.

Kapelko, M., Oude Lansink, A., & Stefanou, S. E. (2016). Investment age and dynamic productivity growth in the Spanish food processing industry. *American Journal of Agricultural Economics*, *98*, 946–961.

Kapelko, M., Oude Lansink, A., & Stefanou, S. E. (2017). The impact of the 2008 financial crisis on dynamic productivity growth of the Spanish food manufacturing industry. An impulse response analysis. *Agricultural Economics*, *48*, 561–571.

Kumbhakar, S. C. (1997). Modeling allocative inefficiency in a translog cost function and cost share equations: An exact relationship. *Journal of Econometrics*, *76*, 351–356.

Kumbhakar, S. C. & Lovell, C. A. K. (2000). *Stochastic Frontier Analysis*. Cambridge: Cambridge University Press.

Kumbhakar, S. C., Parmeter, C. F., & Zelenyuk, V. (2017). Stochastic frontier analysis: Foundations and advances. In S. C. Ray, R. Chambers, & S. C. Kumbhakar (Eds.), *Handbook of Production Economics*, volume 1. Singapore: Springer. forthcoming.

Lambarraa, F., Stefanou, S. E., & Gil, J. M. (2016). The analysis of irreversibility, uncertainty and dynamic technical inefficiency on the investment decision in the Spanish olive sector. *European Review of Agricultural Economics*, *43*, 59–77.

Minviel, J. J. & Sipiläinen, T. (2018). Dynamic stochastic analysis of the farm subsidy-efficiency link: Evidence from France. *Journal of Productivity Analysis*, *50*, 41–54.

Nemoto, J. & Goto, N. (1999). Dynamic data envelopment analysis: Modeling intertemporal behavior of a firm in the presence of productive inefficiencies. *Economics Letters*, *64*, 51–56.

Nemoto, J. & Goto, N. (2003). Measurement of dynamic inefficiency in production: An application of data envelopment analysis to Japanese electric utilities. *Journal of Productivity Analysis*, *19*, 191–210.

O'Donnell, C. J. (2012). Nonparametric estimates of the components of productivity and profitability change in U.S. agriculture. *American Journal of Agricultural Economics*, *94*, 873–890.

O'Donnell, C. J. (2016). Using information about technologies, markets and firm behaviour to decompose a propose productivity index. *Journal of Econometrics*, *190*, 328–340.

Oude Lansink, A., Stefanou, S., & Serra, T. (2015). Primal and dual dynamic Luenberger productivity indicators. *European Journal of Operational Research*, *241*, 555–563.

Parmeter, C. F. & Kumbhakar, S. C. (2014). Efficiency analysis: A primer on recent advances. *Foundations and Trends in Econometrics*, *7*, 191–385.

Pindyck, R. S. & Rotemberg, J. T. (1983). Dynamic factor demands and the effects of energy price shocks. *American Economic Review*, *73*, 1066–1079.

Pitt, M. K., dos Santos Silva, R., Giordani, P., & Kohn, R. (2012). On some properties of Markov Chain Monte Carlo simulation methods based on the particle filter. *Journal of Econometrics*, *171*, 134–151.

Plastina, A. & Lence, S. H. (2018). A parametric estimation of total factor productivity and its components in U.S. agriculture. *American Journal of Agricultural Economics*, *100*, 1091–1119.

Ristic, B., Gordon, N., & Arulampalam, S. (2004). *Beyond the Kalman Filter: Particle Filters for Tracking Applications Partical Filters for Tracking Applications*. Norwood, MA: Artech House.

Rungsuriyawiboon, S. & Stefanou, S. (2007). Primal and dual dynamic Luenberger productivity indicators. *Journal of Business and Economic Statistics*, *25*, 226–238.

Schennach, S. M. (2005). Bayesian exponentially tilted empirical likelihood. *Biometrika*, *92*, 31–46.

Sengupta, J. K. (1995). *Dynamics of Data Envelopment Analysis*. Dordrecht: Kluwer Academic Publishers.

Serra, T., Oude Lansink, A., & Stefanou, S. E. (2011). Measurement of dynamic efficiency: A directional distance function parametric approach. *American Journal of Agricultural Economics*, *93*, 756–767.

Silva, E. & Oude Lansink, A. (2013). Dynamic efficiency measurement: A directional distance function approach. Centro de Economia e Finanças da UPorto Working Paper 2013–07.

Silva, E. & Stefanou, S. E. (2003). Nonparametric dynamic production analysis and the theory of cost. *Journal of Productivity Analysis*, *19*, 5–32.

Silva, E. & Stefanou, S. E. (2007). Dynamic efficiency measurement: Theory and application. *American Journal of Agricultural Economics*, *89*, 398–419.

Skevas, I., Emvalomatis, G., & Brümmer, B. (2018). Productivity growth measurement and decomposition under a dynamic inefficiency specification: The case of German dairy farms. *European Journal of Operational Research*, *271*, 250–261.

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

Tsionas, E. G. & Izzeldin, M. (2018). A novel model of costly technical efficiency. *European Journal of Operational Research*, *268*, 653–664.

Wang, M.-H. & Huang, T.-H. (2007). A study on the persistence of Farrell's efficiency measure under a dynamic framework. *European Journal of Operational Research*, *180*, 1302–1316.