JOURNAL OF APPLIED ECONOMETRICS
J. Appl. Econ. **24**: 675–697 (2009)
Published online 2 April 2009 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/jae.1051

FEASIBLE ESTIMATION OF FIRM-SPECIFIC ALLOCATIVE INEFFICIENCY THROUGH BAYESIAN NUMERICAL METHODS

SCOTT E. ATKINSON[a] AND JEFFREY H. DORFMAN[b]*

[a] *Department of Economics, University of Georgia, Athens, GA, USA*
[b] *Department of Agricultural and Applied Economics, University of Georgia, Athens, GA, USA*

SUMMARY

Both the theoretical and empirical literature on the estimation of allocative and technical inefficiency has grown enormously. To minimize aggregation bias, ideally one should estimate firm and input-specific parameters describing allocative inefficiency. However, identifying these parameters has often proven difficult. For a panel of Chilean hydroelectric power plants, we obtain a full set of such parameters using Gibbs sampling, which draws sequentially from conditional generalized method of moments (GMM) estimates obtained via instrumental variables estimation. We find an economically significant range of firm-specific efficiency estimates with differing degrees of precision. The standard GMM approach estimates virtually no allocative inefficiency for industry-wide parameters.

1. INTRODUCTION

The literature on efficiency measurement has been concerned with measuring economic efficiency in terms of technical and allocative efficiency. Technical efficiency measures the actual input usage relative to the minimum input usage for a given set of outputs or the actual outputs relative to the maximum potential outputs for a given set of inputs. The technical efficiency of a firm is measured relative to the most efficient firm, which defines the production, distance, or cost frontier. Allocative efficiency measures how well firms manage the ratios of inputs in order to minimize the cost of producing a given output. From the first-order conditions for cost minimization, firms must equate the ratio of marginal products to the ratio of input prices.

Allocative and technical efficiency have been measured using either an error-components approach or a fixed-effects approach. With the latter approach, one estimates parameters that measure allocative efficiency in a shadow distance system context by scaling input quantities or in a shadow cost system context by scaling input prices, yielding shadow input quantities and shadow input prices, respectively. Details are provided in Atkinson and Primont (2002). With the error-components approach, maximum likelihood techniques are typically used to estimate the parameters that define the distribution of a two-component error term, where a two-sided error represents noise and a one-sided error represents inefficiency specific to the firm or individual. Firm-specific measures of allocative and technical inefficiency are then obtained from these estimated components. For a survey of this approach see Greene (1997). One major

* Correspondence to: Jeffrey H. Dorfman, Department of Agricultural and Applied Economics, University of Georgia, Athens, GA 30602, USA. E-mail: jdorfman@agecon.uga.edu

drawback to this technique is that incorrect distributional assumptions on the error components will produce inconsistent estimators. A second shortcoming is that, owing to the complexities of using instruments in this setting, researchers typically make the often implausible assumption that the error terms are uncorrelated with the regressors, which must hold in order to avoid endogeneity. One error-components modeling approach has made the one-sided component a function of explanatory variables.[1] If this approach is taken, either these variables must be uncorrelated with the explanatory variables of the model itself to avoid omitted variable bias or appropriate instruments must be employed. Further, results may be very sensitive to the assumed functional form. If the distribution of the composed error is misspecified or if endogeneity is ignored, parameter estimators will be inconsistent.

The fixed-effects approach avoids the need to specify the distribution of the errors or assume the exogeneity of the regressors. This approach easily allows testing for endogeneity and correcting for its presence using instruments. Further, using this approach we can easily obtain robust estimated standard errors. The simplest fixed-effects approach would be to estimate time-invariant, input-specific allocative efficiency parameters. One can still estimate the effect on input over- or under-utilization at the firm level by computing ratios of fitted input quantities. However, the assumption that all firms share common time-invariant allocative efficiency parameters for each input is a priori implausible. Estimation of time-varying, firm-specific (or plant-specific if data are available at this level) allocative efficiency parameters is generally preferred in order to reduce aggregation bias.

Estimation of a full set of firm-specific, time-varying allocative efficiency parameters has never been previously accomplished. Rarely, in fact, have researchers been able to estimate a full set of firm-specific, time-invariant allocative efficiency parameters. Even with panel data and models that are highly nonlinear in the parameters, which should assist identification, the data do not typically contain enough independent variation to identify the full set of parameters measuring allocative efficiency along with the structural parameters that define stochastic cost, revenue, profit, distance, or production frontiers (the standard dual frontier paradigms). A random-effects Bayesian approach has been taken by Kumbhakar and Tsionas (2005), who estimate a translog cost system using Gibbs sampling methods. However, their approach is subject to all the limitations of the classical random-effects estimation approach and requires a full parametric specification of the structural model and the likelihood function. Further, they obtain firm-specific measures of allocative efficiency which are not time-varying.

In this paper, we present a solution to this problem by substituting a semi-parametric Bayesian approach implemented using a Markov chain Monte Carlo (MCMC) technique. Our MCMC approach follows and extends Zellner and Tobias (2001), Kim (2002), and Atkinson and Dorfman (2005) by employing a limited information likelihood function that minimizes the assumptions required for estimation, making it essentially equivalent to Bayesian generalized method of moments (GMM) with instruments. While we are fully parametric in the specification of the structural model, we are non-parametric in the specification of the error distribution. A parallel semi-parametric approach taken by Griffin and Steel (2004), Koop and Tobias (2006), and Koop and Poirier (2004) reverses the choice of the parametric and non-parametric components. They employ a parametric likelihood function with a partially or fully non-parametric structural model.

For a panel of 12 Chilean hydroelectric power plants, we jointly estimate a translog input distance function and the first-order conditions from the dual shadow-cost minimization model. We obtain

[1] See Huang and Liu (1994).

DOI: 10.1002/jae

the posterior densities for plant-specific and time-varying estimates of allocative efficiency, while correcting for any endogeneity of the regressors. We are able to identify and precisely estimate this rich parameterization, when we could not using the classical approach, because we iteratively draw from a series of conditional posterior densities for the parameters using MCMC. Time-varying and plant-specific measures of technical efficiency (TE) are computed residually along with estimates of productivity change (PC), which can be decomposed into technical change (TC) and efficiency change (EC).

Our empirical findings are that energy is under-utilized in five plants (due to limited water availability or restrictions on water release) by as much as 18% and over-utilized in the rest by as much as 18%. Labor is over-utilized in seven plants by as much as 27% and under-utilized in six plants by as much as 14%. Average rates across all plants are considerably smaller. Little movement toward allocative efficiency is observed over time for any plant. Technical efficiency is low for many plants and productivity change and efficiency change vary widely across plants. For example, PC ranges from 2% to 12% across plants for year 9 of the sample. Finally, we re-estimated our model using industry-wide allocative inefficiency parameters in place of plant-specific ones. The aggregate allocative inefficiency parameters indicate no appreciable allocative inefficiency overall and mask important differences among the plant-specific parametric estimates with regard to central tendency as well as precision.

2. FIRM-SPECIFIC ALLOCATIVE AND TECHNICAL INEFFICIENCY

To model firm-specific allocative inefficiency we follow Atkinson and Primont (2002) and employ an input distance function in a cost-minimizing framework to derive a set of estimating equations. We generalize their set-up slightly to add the firm-specific inefficiency measures. To begin, an input distance function is defined as

$$D(\mathbf{y}_t, \mathbf{x}_t) = \sup_{\lambda}\{\lambda : (\mathbf{x}_t/\lambda) \in L(\mathbf{y}_t)\}, \tag{1}$$

where $\mathbf{y}_t$ is an $M \times 1$ vector of outputs, $\mathbf{x}_t$ is an $N \times 1$ vector of inputs, $L(\mathbf{y}_t)$ is the input requirement set, and $\lambda \geq 1$. Its inverse is the measure of technical inefficiency. Then, we assume the typical firm solves the following cost minimization problem:

$$C(\mathbf{y}_t, \mathbf{p}_t) = \min_{\mathbf{x}_t}\{\mathbf{p}_t\mathbf{x}_t : D(\mathbf{y}_t, \mathbf{x}_t) \geq 1\}, \tag{2}$$

where $\mathbf{p}_t$ is a $1 \times N$ vector of input prices. Due to constraints on the optimization process, in most cases, the observed value of $\mathbf{x}_t$ will fail to solve (2). Denote the 'shadow' input quantities that *do* solve (2) by $\mathbf{x}_t^* = [k_{1t}x_{1t}, \ldots, k_{Nt}x_{Nt}]$, where the k_{nt} are measures of input-specific departures of shadow input quantities from actual input quantities and are parameters to be estimated. Subject to a normalization for one input, these measures of allocative inefficiency will be estimated for each plant, but plant-identifying subscripts are suppressed here for simplicity.

The first-order conditions corresponding to (2) are

$$p_{nt} = \mu\frac{\partial D(\mathbf{y}_t, \mathbf{x}_t^*)}{\partial x_{nt}}, \quad n = 1, \ldots, N, \tag{3}$$

DOI: 10.1002/jae

where μ is the Lagrangian multiplier and the derivative is evaluated at the shadow values of the inputs.

Two properties of distance functions can be used to transform equation (3) into an estimable equation. First, the distance function is linearly homogeneous in the inputs, which implies that $\sum_{n=1}^{N} \frac{\partial D(\mathbf{y}_t, \mathbf{x}_t^*)}{\partial x_{nt}} x_n^* = D(\mathbf{y}_t, \mathbf{x}_t^*)$ by Euler's theorem. Second, $D(\mathbf{y}_t, \mathbf{x}_t^*) = 1$ by definition. Thus, if we multiply both sides of (3) by optimal input levels $\mathbf{x}_t^*$, sum over all N inputs, and apply these two properties, we obtain

$$p_n = (\mathbf{p}_t \mathbf{x}_t^*) \frac{\partial D(\mathbf{y}_t, \mathbf{x}_t^*)}{\partial x_{nt}}, \quad n = 1, \ldots, N. \tag{4}$$

The above set of equations along with the distance function comprise the set of equations we employ to estimate firm-specific technical and allocative efficiency.

2.1. Model Specification

To make our econometric model specific to panel data, the shadow input distance function for firm or plant f and time t can be written as

$$1 = D(\mathbf{y}_{ft}, \mathbf{x}_{ft}^*, t) g(\varepsilon_{ft}), \tag{5}$$

where ε_{ft} is a random error. We adopt the translog functional form for the distance function, (5):

$$\begin{aligned}
0 &= \ln D(\mathbf{y}_{ft}, \mathbf{x}_{ft}^*) + \ln g(\varepsilon_{ft}) \\
&= -\beta_{f0} d_f + \sum_m \gamma_m \ln y_{mft} + .5 \sum_m \sum_r \gamma_{mr} \ln(y_{mft}) \ln(y_{rft}) \\
&+ \sum_m \sum_n \gamma_{mn} \ln y_{mft} \ln x_{nft}^* + \sum_n \gamma_n \ln x_{nft}^* \\
&+ .5 \sum_n \sum_l \gamma_{nl} \ln x_{nft}^* \ln x_{lft}^* + \sum_m \gamma_{mt} \ln y_{mft} w \\
&+ \sum_n \gamma_{nw} \ln x_{nft}^* w + \gamma_w w + .5 \gamma_{ww} w^2 \\
&+ \gamma_{wt} wt + \gamma_{t1} t + .5 \gamma_{t2} t^2 + \ln g(\varepsilon_{ft})
\end{aligned} \tag{6}$$

where

$$g(\varepsilon_{ft}) = \exp(\upsilon_{ft} - u_{ft}^*). \tag{7}$$

The two-sided disturbance is υ_{ft}, while the one-sided component is

$$u_{ft}^* = u_{ft} - \beta_{f0} d_f \geq 0 \tag{8}$$

which captures technical inefficiency. The variable w measures relative hydrologic conditions (an important exogenous shifter of the production function), while d_f are plant dummies. As proposed

DOI: 10.1002/jae

by Cornwell *et al.* (1990), we specify u_{ft} in terms of plant-specific dummies and these dummies interacted with linear and quadratic trends:

$$u_{ft} = \beta_0 + \beta_{f0}d_f + \beta_{f1}d_f t + \beta_{f2}d_f t^2. \tag{9}$$

Given (9), our definition of u^*_{ft} in (8) accounts for the fact that we subtracted $\beta_{f0}d_f$ from u_{ft} and added its negative to (6) in order to obtain a fixed-effects panel data estimator that eliminates unobserved time-invariant heterogeneity. This adjustment is used below in the calculation of technical efficiency.

In this paper, we are more general than any previous approach with respect to modeling allocative efficiency across firms and time. Since we utilize plant-level panel data, we specify a plant-specific, time-varying allocative efficiency measure (k_{nft}), for each of N inputs, as a function of time and time squared:

$$k_{nft} = \exp(\kappa_{nf} + \kappa_{nf1}t + \kappa_{nf2}t^2), \tag{10}$$

which allows for allocative efficiency to change non-monotonically over time. In comparison, previous researchers have imposed various restrictions on this general model, limiting flexibility across firms and time. Atkinson and Halvorsen (1984) used a cross-section of firms to estimate a shadow cost system, comprised of a shadow cost equation and a set of derived share equations. The use of cross-sectional data required restricting k_{nft} to

$$k_n = \exp(\kappa_n), \tag{11}$$

which allows estimation of only input-specific, time-invariant allocative efficiency parameters. Atkinson and Cornwell (1994) used panel data on firms to estimate a shadow cost system. In theory, panel data should allow unrestricted estimation of (10). However, they were unable to identify the time-varying component of k_{nft}, which was therefore restricted to

$$k_{nf} = \exp(\kappa_{nf}), \tag{12}$$

which allows for firm-specific allocative efficiency parameters which are time-invariant. Atkinson *et al.* (2003) utilize panel data on firms to estimate a shadow distance system, comprised of a distance equation as a function of shadow quantities and associated price equations. Again, due to problems with identifying the full set of parameters in (10), they restrict k_{nft} to

$$k_{nft} = \exp(\kappa_{nf} + \kappa_{n1}t + \kappa_{n2}t^2), \tag{13}$$

which allows for firm-specific, time-invariant parameters but only industry-wide time-varying parameters (κ_{n1} and κ_{n2}) that are shared across firms.

Identification of the flexible model in (10) still requires a restriction on the κ_{nft}. Owing to linear homogeneity of the distance function in input quantities, we must normalize κ_{nft} for some input n for each firm f. This implies that we can only measure the over- or under-utilization of one input relative to another; thus, we set $\kappa_{nft} = 1$ for a numeraire input $\forall t, f$. The specific choice of the numeraire does not affect any model parameters other than the absolute values of the κ_{nft}; however, their relative values remain unchanged, regardless of the choice of the numeraire.

DOI: 10.1002/jae

We avoid what Bauer (1990) called the 'Greene problem', first noted by Greene (1980), as the problem or difficulty in specifying estimable models that allow allocative and technical inefficiency to be correlated. The Greene problem has become a search for a translog (or more generally a flexible functional form) specification that incorporates correlated measures of allocative and technical inefficiency. Bauer (1990) summarizes considerable research devoted to solving the Greene problem using error components models which are estimated via maximum likelihood. We avoid both the misspecification problems inherent with this approach and solve the Greene problem by measuring allocative efficiency using parameters which scale input quantities. This approach allows correlation between the parameters of the distance function (which include parameters measuring allocative efficiency) and the one-sided residual, $\hat{u}_{ft}$, from which we derive our measure of technical efficiency.

We substitute the restrictions in (7), (9), and (10), along with those that impose symmetry and linear homogeneity (see Atkinson and Primont, 2002, for details) into the stochastic translog shadow distance system (6). Taking derivatives of (6), we can specify (4) in terms of the distance function parameters as

$$\begin{aligned}
p_n = (\mathbf{p}_t\mathbf{x}_t^*)\Bigg\{ & \left[\gamma_n + \sum_m \gamma_{mn}\ln y_{mft} + \sum_l \gamma_{nl}\ln x^*_{lft} + \gamma_{nw}w\right] \times \\
& \exp\Bigg[-\beta_{f0}d_f + \sum_m \gamma_m \ln y_{mft} + 0.5\sum_m\sum_r \gamma_{mr}\ln(y_{mft})\ln(y_{rft}) \\
& + \sum_m\sum_n \gamma_{mn}\ln y_{mft}\ln x^*_{nft} + \sum_n \gamma_n \ln x^*_{nft} \\
& + 0.5\sum_n\sum_l \gamma_{nl}\ln x^*_{nft}\ln x^*_{lft} + \sum_m \gamma_{mt}\ln y_{mft}w \\
& + \sum_n \gamma_{nw}\ln x^*_{nft}w + \gamma_w w + 0.5\gamma_{ww}w^2 \\
& + \gamma_{wt}wt + \gamma_{t1}t + 0.5\gamma_{t2}t^2\Bigg] \times \\
& \Bigg[-\beta_{f0}d_f + \sum_m \gamma_m \ln y_{mft} + 0.5\sum_m\sum_r \gamma_{mw}\ln(y_{mft})\ln(y_{rft}) \\
& + \sum_m\sum_n \gamma_{mn}\ln y_{mft}\ln x^*_{nft} + \sum_n \gamma_n \ln x^*_{nft} \\
& + 0.5\sum_n\sum_l \gamma_{nl}\ln x^*_{nft}\ln x^*_{lft} + \sum_m \gamma_{mt}\ln y_{mft}w \\
& + \sum_n \gamma_{nw}\ln x^*_{nft}w + \gamma_w w + 0.5\gamma_{ww}w^2 \\
& + \gamma_{wt}wt + \gamma_{t1}t + 0.5\gamma_{t2}t^2\Bigg]^{-1} \\
& \times \left[\frac{1}{x^*_{nft}}\right]\Bigg\}, n = 1, \ldots, N. \qquad (14)
\end{aligned}$$

DOI: 10.1002/jae

We then append a random error term to the shadow distance function and N derived price equations in (14) to obtain a system of $(N+1)$ nonlinear equations with multiple cross-equation restrictions which we refer to as the shadow distance system. Note that the high degree of nonlinearity in the unknown parameters in equations (14) should aid identification of the allocative efficiency parameters.

2.2. Measurement of Allocative Inefficiency

By taking ratios of equations in (4), we obtain the conditions for cost minimization in terms of shadow quantities and actual prices. For firm f at time t, we can directly estimate relative over- and under-utilization of any pair of inputs, x_{nft} and x_{lft}, in comparison to the cost-minimizing ratio, $(k_{nft}x_{nft})/(k_{lft}x_{lft})$, by computing $\hat{k}_{nft}/\hat{k}_{lft}$. We argue that frequently researchers and policy makers are more interested in shadow quantities than shadow prices, which give a fundamental advantage to the shadow distance system over the shadow cost system. Examples include the effects of quotas or restrictive work rules on input usage, inefficient input usage due to rate of return regulation, and the impact of government subsidies or tariffs in agriculture on input usage.

2.3. Measurement of Technical Efficiency

Following the estimation of (6), we compute levels of TE, EC, TC, and PC. To compute TE, after estimating (6), we calculate the negative of $-\hat{\beta}_{f0}d_f + \hat{\upsilon}_{ft} - \hat{u}^*_{ft}$ which equals $\hat{\beta}_{f0}d_f - \hat{\upsilon}_{ft} + \hat{u}^*_{ft} = \hat{u}_{ft} - \hat{\upsilon}_{ft}$. These fitted values are then regressed on the right-hand side of (9), where we must impose a restriction on β_0 or β_{f0} for one f. The fitted value of this regression is an estimator of u_{ft}. Since the non-negativity of the u_{ft} is not imposed in estimation, we now impose this condition by computing $\hat{u}_t = \min_f(\hat{u}_{ft})$ and then constructing $\hat{u}^{\mathcal{F}}_{ft} = \hat{u}_{ft} - \hat{u}_t \geq 0$. We create the standard estimate of plant f's level of technical efficiency in period t, TE_{ft}, as

$$\mathrm{TE}_{ft} = \exp(-\hat{u}^{\mathcal{F}}_{ft}), \tag{15}$$

where our normalization of $\hat{u}^{\mathcal{F}}_{ft}$ guarantees that $0 \leq \mathrm{TE}_{ft} \leq 1$. This measure is relative to the 'best practice' plant. Our specification of TE solves the Greene problem, since the measure of technical efficiency in (15) is a function of the estimated residuals from (6), which in turn are functions of the estimated allocative efficiency parameters. Hence, measures of allocative and technical efficiency are correlated. Given the estimates of TE_{ft} obtained from (15), we then calculate EC_{ft}, the rate at which a firm is approaching the isoquant, as the change in technical efficiency:

$$\mathrm{EC}_{ft} = \Delta\mathrm{TE}_{ft} = \mathrm{TE}_{ft} - \mathrm{TE}_{f,t-1}. \tag{16}$$

We measure TC, the movement inward of isoquants, as a discrete approximation which involves computing the difference between the estimated frontier distance function in periods t and $t-1$ holding output and input quantities constant:

$$\begin{aligned}\mathrm{TC}_{ft} &= \ln \hat{D}^*(\mathbf{y}_t, \mathbf{x}^*_t, t) - \ln \hat{D}^*(\mathbf{y}_t, \mathbf{x}^*_t, t-1) \\ &= \ln \hat{D}(\mathbf{y}_t, \mathbf{x}^*_t, t) - \ln \hat{D}(\mathbf{y}_t, \mathbf{x}^*_t, t-1) + (\hat{u}_{t-1} - \hat{u}_t).\end{aligned} \tag{17}$$

DOI: 10.1002/jae

Thus the change in the frontier intercept, $\hat{u}_t$, affects TC as well as EC. Finally, given EC and TC, we construct estimates of productivity change, PC, as

$$\mathrm{PC}_{ft} = \mathrm{TC}_{ft} + \mathrm{EC}_{ft}. \tag{18}$$

One should note that these definitions of PC, TC, and EC produce the same percentage rates of increase as one would obtain using the Koop *et al.* (1999) measure of $\mathrm{PC}_{ft} = \mathrm{TC}_{ft}\mathrm{EC}_{ft}$, where all measures are based on index numbers. Using our notation, their measure of TC is an index number defined as $\hat{D}(\mathbf{y}_t, \mathbf{x}_t^*, t)/\hat{D}(\mathbf{y}_t, \mathbf{x}_t^*, t+1)$. If, for example there is positive TC equal to 1.03, our estimates are equal to $\partial \ln \hat{D}(\mathbf{y}_t, \mathbf{x}_t^*, t)/\partial t = \ln \hat{D}^*(\mathbf{y}_t, \mathbf{x}_t^*, t) - \ln \hat{D}^*(\mathbf{y}_t, \mathbf{x}_t^*, t-1)$, so that TC would be simply 0.03. Both measures imply a 3% growth in TC. The same equivalence holds true for EC and hence PC.

3. ESTIMATION

3.1. An Error Components Approach

One possible estimation procedure would be to completely specify the full likelihood and estimate an error-components model in a Bayesian framework. Zellner *et al.* (1988) point out the difficulties in accurate specification of the full likelihood. In addition, if the composed error term is correlated with the explanatory variables one must employ instrumental variables. We argue that this is far easier to accomplish using a Bayesian GMM approach than a full maximum-likelihood one. In addition, by using a Bayesian GMM approach we obtain an estimator of the covariance matrix that is consistent subject to autocorrelation and heteroskedasticity of unknown form. In terms of our distance system, a typical composed-error specification is that

$$0 = \ln D(\mathbf{y}_{ft}, \mathbf{x}_{ft}^*, t) + \ln g(\varepsilon_{ft}) = \ln D(\mathbf{y}_{ft}, \mathbf{x}_{ft}^*, t) + \upsilon_{ft} - u_{ft} \tag{19}$$

and

$$p_n = (\mathrm{p}_t \mathbf{x}_t^*) \frac{\partial D(\mathbf{y}_t, \mathbf{x}_t^*)}{\partial x_{nt}} + \eta_{ft}, \quad n = 1, \ldots, N. \tag{20}$$

Various assumptions about the independence of the error terms and their joint distribution in these two equations have been made and each may be incorrect, resulting in model misspecification. See Greene (1997) for a good summary. In addition, one must assume that the error terms are uncorrelated with the regressors. Estimated parameters will be inconsistent if regressors are endogenous and dealing with endogeneity within this context is problematic.

3.2. An Alternative Allowing Endogeneity

To derive estimators for the shadow input distance function system, we use a limited-information Bayesian system estimator to avoid the difficulties in accurately specifying the full likelihood. In our estimation approach, we follow and generalize Kim (2002) and Zellner and Tobias (2001).[2]

[2] Zellner and Highfield (1988) and Zellner (1998) develop early Bayesian method of moments (BMOM) estimators that clearly presage Kim's approach. Other applications of BMOM can be found in Green and Strawderman (1996) and LaFrance (1999).

DOI: 10.1002/jae

Kim proves that maximizing entropy subject to a restriction on a GMM criterion function yields an optimal limited-information likelihood function (LILF). We advance Kim's approach by treating the covariance of the errors as well as unknown parameters of our distance system as random variables and constructing a joint LILF. Interested readers can see Atkinson and Dorfman (2005) for more details.

The form of the LILF and the resulting posterior distribution depend on the moment conditions that serve as the basis for the criterion function. Following the development in Kim (2002), we start with a standard first moment condition for the parameter vector γ:

$$E[h(\gamma|\Omega, D)] = 0 \tag{21}$$

where γ is a vector of regression model parameters, $h(\gamma)$ is a vector of the errors, Ω is the covariance matrix of the regression model's stochastic error terms, and D represents the data, including instruments. This restriction sets the conditional expectation of the error terms equal to zero. A useful second moment condition on γ is

$$E[h(\gamma)h(\gamma)'|D] = S \tag{22}$$

where S is the standard GMM estimator of the covariance matrix of $h(\gamma)$. Other second moment conditions are possible, but this one leads to a rather convenient LILF at no particular cost in terms of loss of flexibility in estimation. Recall that we are deriving a Bayesian GMM estimator, so increasing similarities between traditional GMM estimators and this new one makes sense.

We now extend Kim (2002) to include the random covariance matrix parameters in our LILF by adding two moment conditions on Ω:

$$E[\mathrm{tr}(\Xi\Omega^{-1})|D] = \xi \quad \text{and} \quad E[\ln|\Omega||D] = \tau \tag{23}$$

where Ξ is the sum of squared residuals matrix, while ξ and τ are scalar constants. These moment conditions impose enough regularity on the otherwise unrestricted distribution of Ω to result in a conditional limited information posterior (LIP) distribution for Ω in the form of an inverted Wishart distribution; see Zellner and Tobias (2001) for details of the univariate case which they pioneered.

Any selection from the set of admissible LILFs is obviously somewhat ad hoc, but we defend our selection process as being in the spirit of GMM estimators. Thus, we choose the least informative (most diffuse) LILF from the admissible set in order to impose the minimum amount of assumptions on the estimation procedure.

From the set of admissible functions $\mathcal{F}$ that satisfies the above moment conditions, the least informative LILF, f, is found by solving the optimization problem

$$\mathrm{argmax}_{f\varepsilon\mathcal{F}} - \int f(D|\gamma, \Omega)\ln f(D|\gamma, \Omega)\mathrm{d}\gamma\mathrm{d}\Omega. \tag{24}$$

The solution is

$$\hat{f}(D|\gamma, \Omega) = c_o|\Omega|^{-c_1}\exp\left[-c_2 h(\gamma)'S^{-1}h(\gamma) - c_3\mathrm{tr}(\Xi\Omega^{-1})\right]. \tag{25}$$

DOI: 10.1002/jae

Recalling that $h(\gamma)$ is the vector of residuals from the model, inspection of the above LILF shows $\hat{f}$ to be the product of a distribution from the exponential family for γ and an inverted Wishart with respect to Ω, where c_o, c_1, c_2, and c_3 are constants. If one used the LILF in (25) as a likelihood function and found the values of γ and Ω which maximize it, the result is the standard GMM estimator. The proof of this follows easily from that in Kim (2002, equation 3.8) for the case with a known Ω.

The prior distribution used in our application is a product of independent priors on the structural parameters of the distance function, the prior on the covariance matrix of the vector of errors, and a set of indicator functions that restrict prior support to the region where the theoretical restrictions from economic theory are satisfied. This prior distribution can be written as

$$p(\gamma, \Omega) \propto \mathrm{MVN}(g_o, H_o)|\Omega|^{-(m+1)/2}\mathrm{I}(\gamma, \mathcal{R}) \tag{26}$$

where MVN is the multivariate normal distribution, g_o is the vector of prior means of the parameters in γ, H_o is the prior variance–covariance matrix of these same parameters, $I(\gamma, \mathcal{R})$ represents the indicator function that equals one when the restrictions are satisfied and zero otherwise, and m is the number of equations in our system. Our actual priors reflect our basic ignorance of the production function parameters and the expectation of complete allocative efficiency.[3]

The indicator function part of the prior restricts positive prior (and posterior) support to the region, $\mathcal{R}$, that satisfies monotonicity for all inputs and for the output. Ideally, monotonicity would be satisfied at 100% of our data points. However, we allow for potential measurement errors by requiring that monotonicity for inputs and the output be satisfied for 95% and 99% of our observations, respectively.[4]

Having derived the LILF and defined the prior density $p(\gamma, \Omega)$, we apply Bayes theorem using the LILF in place of a standard likelihood function, and derive a limited-information posterior distribution:

$$p(\gamma, \Omega|D) = p(\gamma, \Omega)\hat{f}(D|\gamma, \Omega)\mathrm{c}^{-1}, \tag{27}$$

where c is the normalizing constant. Only values or expressions for the two constants, ξ and τ, in the moment conditions for Ω from equation (23) are needed to derive the precise limited-information posterior. Careful choice of these two constants leads to a limited-information posterior with 'standard' parameters:

$$p(\gamma, \Omega|D) = c_o|\Omega|^{-(n-k+m+1)/2}\exp\left[-\frac{1}{2}(\gamma-\gamma_p)'\Psi_p^{-1}(\gamma-\gamma_p) - \frac{1}{2}\mathrm{tr}(\Xi\Omega^{-1})\right]\mathrm{I}(\gamma, \mathcal{R}), \tag{28}$$

where

$$\gamma_p = \Psi_p(H_o^{-1}g_o + \Psi_m^{-1}\gamma_m) \tag{29}$$

[3] The vector g_o is set to zero. The matrix H_o is a diagonal covariance matrix, whose elements are 1.0×10^3 for the basic structural parameters of the distance function, 1.0×10^{-4} for the time-invariant, plant-specific allocative parameters, 1.0×10^{-5} for the parameters on time interacted with the plant-specific allocative parameters, and 1.0×10^{-9} for the parameters on time-squared interacted with the plant-specific allocative parameters. Without such priors, the terms involving t and t^2 made the Markov chain wander to unreasonably large or small parameter draws without recovering.

[4] See Lim and Shumway (1992) for a similar recognition that not all data points will satisfy monotonicity due to measurement error.

DOI: 10.1002/jae

$$\Psi_p = (H_o^{-1} + \Psi_m^{-1})^{-1}, \tag{30}$$

(where, to reiterate, g_o is the vector of prior means of the parameters in γ and H_o is the prior variance–covariance matrix of these parameters, while γ_m is the standard GMM estimator of γ and Ψ_m is the standard GMM estimator of the covariance matrix of γ_m, given the set of identifying restrictions we have specified).[5] The limited-information posterior distribution in (28) is a truncated version of the standard multivariate normal-inverted Wishart distribution common in Bayesian econometrics, although with a nonstandard mode (due to the replacement of the usual ML estimator with a GMM estimator in (29) and (30)).

It is worth noting that because the posterior distribution is derived by combining the prior and the limited information likelihood, potential point estimators such as the posterior mean will not exactly satisfy the moment conditions of equations (21) and (22). However, even in non-Bayesian method of moments or GMM estimation, when overidentifying restrictions are employed, the resulting estimators do not exactly satisfy every moment condition. The point of such GMM estimators, both standard and Bayesian ones, is to find estimators that are 'close' to meeting the moment conditions by balancing deviations from exact equality among the different restrictions. This is no different from a seemingly unrelated estimator which balances residuals across equations to find the best estimator according to an overall objective function designed to minimize the sum of weighted squared residuals.

The posterior distribution defined above in (28) does not allow for analytical calculation of posterior means and medians of the parameters in γ or functions of those parameters (such as the k_{nft} that we are most interested in). Therefore, numerical methods must be employed to approximate the posterior distribution and estimate the posterior means and medians that are of interest. For this purpose, we used an MCMC approach, specifically Gibbs sampling with an accept–reject step for the imposition of the monotonicity condition. For more details on Gibbs sampling and MCMC methods, see Casella and George (1992) or Tierney (1994).

The Gibbs sampler consists of repeated draws from conditional distributions of subsets of parameters from which we can more easily generate random draws than from the full posterior. In this application, the conditional distributions are a truncated MVN for γ conditional on Ω and an inverted Wishart for Ω conditional on γ. The conditional posterior for γ is further broken down into the subsets $(\gamma_d, \gamma_s, \kappa_E, \kappa_L)$, where γ_d is a vector of plant dummies, γ_s is a vector of the other structural coefficients, and $\kappa_n = \{\kappa_{nf}, \kappa_{nf1}, \kappa_{nf2}\}$, $n = L, E$, is a vector of parameters measuring allocative inefficiency. Draws for each subset of parameters are accomplished sequentially, conditional on the covariance matrix and the other elements of γ. Draws from the truncated MVN were accomplished by drawing from the untruncated distribution and discarding draws that were not within the region R.

3.3. Comparisons with the Exact Likelihood Approach

To see how the estimation technique employed here varies from the Bayesian technique presented in Kumbhakar and Tsionas (2005), Table I presents a side-by-side comparison of the MCMC

[5] The constants ξ and τ could be estimated but that would greatly complicate the estimation algorithm by adding a numerical optimization step requiring a quasi-Newton or equivalent search algorithm for each loop through the Gibbs sampler that will be employed to approximate the posterior. In return, one would get a more accurate marginal posterior distribution for Ω. We choose to simply set the two constants to convenient values instead, given the interest is on a subset of the parameters in γ.

DOI: 10.1002/jae

algorithms used by each approach. Kumbhakar and Tsionas employ a fully parametric model and likelihood function which leads to a more complex model even though it is more restrictive in its inability to include heteroskedasticity, autocorrelation, and endogeneity without further extensions. Our approach differs in its reliance on a limited information likelihood function which can incorporate heteroskedasticity, autocorrelation, endogeneity, and the use of instruments easily. Further, our approach is implemented without needing to draw from nonstandard distributions, use importance sampling, or rely on Metropolis–Hastings steps in our Gibbs sampler. Thus, while we do condition on two minor parameters of the limited information likelihood function, our method is based on standard distributions and statistical functions built into many popular software packages. Further, we believe that a much larger percentage of applied econometricians who are not Bayesian specialists could successfully estimate a model using our approach than could employ the more complex Kumbhakar and Tsionas methodology.

The steps in Table I outline the Gibbs sampler used in each approach to approximate the posterior distributions of the unknown parameters. In both cases, a draw is made from a conditional

Table I. Comparison of Methodologies

Atkinson–Dorfman	Kumbhakar–Tsionas
1. Structural parameters	1. Structural parameters
a. Estimate γ_m and then compute $\gamma_p \mid \kappa, \Omega, D$ using (29). Obtain Ψ_m and then compute $\Psi_p \mid \kappa, \Omega, D$ using (30).	a. Designate previous values of β by $\beta^{(i-1)}$
b. Draw $\gamma \sim MVN(\gamma_p, \Psi_p) \mid \kappa, \Omega, D$.	b. $\beta \sim f_\beta(\cdot)$, which cannot be sampled.
c. Truncate $f(\gamma)$ to impose monotonicity using accept-reject algorithm.	c. Draw $\tilde{\beta} \sim g\beta(\cdot) = \mathrm{MVN}(\overline{\beta}^{(i-1)}, \overline{V}^{(i-1)}) \mid \xi, \Omega, \sigma_u, u, \Sigma, D$
	d. Draw $a \sim U[0, 1]$.
	e. Accept draw from $g\beta(\cdot)$ if $a < \min\left\{1, \dfrac{f_\beta(\tilde{\beta})/g_\beta(\tilde{\beta})}{f_\beta(\beta^{(i-1)})/g_\beta(\beta^{(i-1)})}\right\}$.
	f. Else, $\beta^{(i)} = \beta^{(i-1)}$.
2. Inefficiency parameters	2. Inefficiency parameters
a. Estimate $\kappa_m \mid \gamma, \Omega, D$; obtain κ_p, $\Psi_{p,\kappa}$.	a. Designate previous value of ξ by $\xi^{(i-1)}$.
b. Draw $\kappa \sim MVN(\kappa_p, \Psi_{p,\kappa}) \mid \gamma, \Omega, D$.	b. $\xi \sim f_\xi(\cdot)$, which cannot be sampled.
	c. Compute $\overline{\Omega}$ and draw $\tilde{\xi} \sim g_\xi(\cdot) = \mathrm{MVN}(0, \overline{\Omega}) \mid \beta, u, \sigma_u, \Omega, \Sigma, D$.
	d. Draw $a_k \sim U[0, 1]$.
	e. Accept draw from $g_\xi(\cdot)$ if $a_k < \min\left\{1, \dfrac{f_\xi(\tilde{\xi})/g_\xi(\tilde{\xi})}{f_\xi(\xi^{(i-1)})/g_\xi(\xi^{(i-1)})}\right\}$.
	f. Else, $\xi^{(i)} = \xi^{(i-1)}$.
c. Compute $k = \exp(\eta'\kappa)$, where $\eta' = (1, t, t^2)$.	g. Compute $k = \exp(\xi)$.
d. Compute u as a residual	h. Obtain $\overline{\sigma}_u$ and draw $u \sim trN(0, \overline{\sigma}_u) \mid \beta, \xi, \Omega, \sigma_u^2, \Sigma, D$.
e. Compute technical inefficiency from $u \mid \gamma, \kappa$, and D using (15).	i. Compute technical inefficiency from u similarly.
3. Covariance matrix	3. Covariance matrix
a. Draw $\Omega \sim IW \mid \gamma, \kappa, D$.	a. Draw $\Sigma^{-1} \sim W \mid \beta, \xi, u, \sigma_u, \Omega^{-1}, D$.
	b. Draw $\Omega^{-1} \sim W \mid \beta, \xi, \sigma_u, \Sigma^{-1}, D$.
	c. Draw $\dfrac{\underline{q} + u'u}{2\sigma_u^2} \sim \chi^2_{n+\underline{n}} \mid \beta, \xi, u, \Omega^{-1}, \Sigma^{-1}, D$.
4. Return to 1a.	4. Return to 1a.

DOI: 10.1002/jae

distribution in each step. When the end of the list of steps is reached, one returns to step 1 and repeats. The draws form an empirical approximation to the joint posterior distribution.

For the Atkinson–Dorfman approach, the steps are slightly simplified in places. Since we define $\kappa = \{\kappa_{Lf}, \kappa_{L1t}, \kappa_{L2t}, \kappa_{Ef}, \kappa_{E1t}, \kappa_{E2t}\}$, reference to drawing κ in Table I is actually a reference to drawing $\kappa_{Lf}, \kappa_{L1t}, \kappa_{L2t}, \kappa_{Ef}, \kappa_{E1t}$, and κ_{E2t} in six consecutive steps. Since we define $\gamma = (\gamma_d, \gamma_s)$, reference to drawing γ is actually a reference to drawing γ_d and then γ_s in two consecutive steps. Each consecutive step is conditional on the draws of the previous ones. We reiterate that Ω is the covariance matrix of the error terms, γ_p is the posterior mean of γ and Ψ_p is its posterior covariance matrix. Similarly, κ_p is the posterior mean of κ and $\Psi_{p,\kappa}$ is its posterior covariance matrix.

Using the notation in equation (11) of Kumbhakar and Tsionas (2005), their model is

$$\mathrm{y} = X(\xi)\beta + \phi(\xi, \delta) + \upsilon + \begin{pmatrix} u \\ 0_{(m-1)n} \end{pmatrix} \tag{31}$$

Additionally, they define $\overline{\beta}$ and $\overline{V}$ as the posterior mean and covariance for β, while Ω and Σ represent the covariance matrices for ξ and the error term, υ, respectively. The parameters m and n are the number of share equations and observations, respectively. The scalar σ_u refers to the standard deviation of u, which is distributed as half-normal, while $\underline{n}$ and $\underline{q}$ are hyperparameters from the prior on σ_u. Further, $\overline{\Omega}$ represents a posterior covariance matrix for ξ and $\overline{\sigma}_u$ a posterior variance for u. Note that in place of our k, they employ $\xi = in(k)$. Finally, D represents the data in both approaches.

Based on this table, one can also see the simplicity of our method compared to that of Kumbhakar and Tsionas (2005) both in terms of the number of steps required and the distributional assumptions required to obtain estimates of allocative and technical efficiency. Specifically, we avoid their Metropolis–Hastings sampling in steps 1 and 2.

In step 3, we are required to draw from only one of three distributions that they draw from. Further, their approach requires that they discard the first 15,000 MCMC draws and use only the last 5000. With our approach we quickly overcome starting value problems and keep the vast majority of our draws.

3.4. Implementation of the Gibbs Sampler

A total of 12,000 Gibbs draws were generated from four separate chains. Each chain was 3500 draws long, with the first 500 discarded to remove dependence on initial starting values (standard GMM estimates were used for that purpose). Convergence was checked by confirming the posterior means of the four separate chains were statistically equivalent. For example, the posterior means and medians of the TE measures did not differ by more than 0.5% in any case across the chains.

We impose certain restrictions in order to achieve identification. First, we impose the restriction $\beta_0 = 0$ in (9), in order to identify the coefficients of the other plant-specific dummies, $\beta_{f0}, , \forall f$. Second, we restrict the allocative inefficiency parameters for one input to achieve identification of the others; this is done by setting $k_{nft} = 1, \forall t$ for one n.

Joint estimation of this system using GMM (which is a part of the formula for the posterior mode of our Bayesian estimator) requires that the model satisfy the moment conditions $E(\upsilon_{ft}|\mathbf{z}_{ft}) = 0$, where $\mathbf{z}_{ft}$ is a vector of instruments. Although not strictly Bayesian, we report the standard

DOI: 10.1002/jae

Hansen (1982) J-test of overidentifying restrictions to determine the validity of the instrument set that is used to estimate our distance system using GMM. We fail to reject the null hypothesis that the moment conditions are satisfied when we employ the following instrument set: all plant-level dummies, the interaction of an 'older' vintage variable (pre-1948 = 1) and log output, the log of the real node price of electricity, a variable measuring the relative hydrologic conditions (water level relative to average water level), w, w^2, w^3, w^4, t, t^2, t^3, the interaction of the run-of-river dummy with monthly dummies, the interaction of the logs of inputs and the log of output with time, the interaction of the logs of inputs with w, log output, the interaction of the run-of-river dummy with output, and log output squared. We allow for heteroskedasticity and autocorrelation of unknown form by computing the consistent covariance matrix following Newey and West (1987) with 10 monthly lags. The lag length was chosen in order the minimize the GMM criterion function. Based on the J-test, we easily accept the null hypothesis of the validity of the overidentifying conditions. Because we must perform this test for each draw of our Gibbs sampler, results vary. However, we always find a p-value in excess of 0.50 with 192 degrees of freedom in the step where we free up the structural parameters. We find that the instruments are not weak based on a regression of each endogenous variable on the list of instruments yielding F statistics greater than the rule-of-thumb value of 10.[6]

For the duality between input prices and quantities to be valid, the input shadow distance function must be monotonically increasing in inputs and monotonically decreasing in outputs. Our estimated model satisfies the required monotonicity properties for inputs and outputs for at least 99% and 95% of the data points, respectively (since that condition was imposed in estimation by restricting draws to the region R).

4. DATA AND RESULTS

4.1. Data

Our sample is a rotated and unbalanced panel, consisting of 13 Chilean hydroelectric power generation plants, observed monthly for a maximum of 141 data points per plant spanning April 1986 to December 1997. The monthly frequency is designed to capture the considerable variation in the country's hydrologic conditions throughout the year. Table II lists the 13 plants for which we have observations, their controlling firm, year of initial service, type of hydro generation, and MW capacity. The four controlling firms really constitute three—Endesa, Gener, and Pilmaiquén—since Endesa owns Pehuenche. Thus, our final sample consists of 13 plants (owned by three firms) with a total of 1935 monthly observations. Observations per plant range from 49 to 141.

The panel was unbalanced for two reasons. The first and most important reason is the staggered initial in-rotation of plants following their construction. The second and less important reason is the limited out-rotation by some plants due to plant-specific mechanical problems.

At the plant level, we record the output quantity (Q), the price per unit of output, and the price and quantity of three inputs—labor (L), capital (K), and water (E)—which we also refer to as energy. All the quantities and prices have been normalized by their means before taking their logarithms. Prices are all in real terms. Full details of the dataset can be found in Atkinson and Halabí (2004). We arbitrarily normalize k_{kft} to 1 for all plants and time periods.

[6] A variety of instrument sets were examined. Of these only a few passed the J-test. Among these, only minor differences appeared in the final results.

Table II. Hydroelectric power plants, December 1997

Plant ID	Plant	Utility	Year	Type	MW
1	Alfalfal	Gener	1991	Run-of-river	160
2	Queltehues	Gener	1928	Run-of-river	41
3	Sauzal and Sauzalito	Endesa	1948	Run-of-river	86
4	Rapel	Endesa	1968	Reservoir	350
5	Canutillar	Endesa	1990	Reservoir	145
6	Cipreses	Endesa	1955	Reservoir	101
7	Isla	Endesa	1963	Run-of-river	68
8	Abanico	Endesa	1948	Run-of-river	136
9	Antuco	Endesa	1981	Reservoir	300
10	Pehuenche	Pehuenche	1991	Reservoir	500
11	Curillinque	Pehuenche	1993	Run-of-river	85
12	Pilmaiquén	Pilmaiquén	1944	Run-of-river	39
13	Pullinque	Pilmaiquén	1962	Run-of-river	49

Note: provided by the Economic Load Dispatch Centre (ELDC), Chile.

Total labor utilization has fallen steadily from 1990 and capital investment has risen steadily through 1995. However, water utilization has been far more variable. For example, water usage (and as a result, output) declined dramatically during the years 1994–1996.

4.2. Empirical Results

Although the plant-specific allocative inefficiency measures all vary by year, the empirical estimates are exceedingly stable. Even given the flexible specification of the k_{nft} in equation (10), the estimates are fairly constant over time. Table III presents posterior means, medians, and standard errors for year 12 estimates of each plant's k_E, measured relative to capital. The posterior medians and means are very similar, which reflects the very symmetric nature of almost all the posterior distributions of the k_Es. In fact, plant 3 has the largest discrepancy between mean and median (0.978 versus 0.990) along with the largest posterior standard error (0.069). The reason behind this can be seen in Figure 1, which reveals that the posterior distribution for this plant's k_E is bimodal, with a small second mode around 0.75. Figures 1 and 2 display the posterior distributions for all 13 plant-specific k_E values. The figures make clear that while no plant's k_E is statistically different from all the others, some plants have allocative inefficiency measures that can be statistically distinguished from a number of other plants. Further, most plants are reasonably good at allocative efficiency, with three plants clearly over-utilizing energy relative to capital and two under-utilizing.

Table IV presents the posterior means, medians, and standard errors for year 12 estimates of each plant's k_L, again measured relative to capital. The posterior medians and means are even more similar than for energy with no bimodality present. Figures 3 and 4 display the posterior distributions for all 13 plant-specific k_L values. For labor, we find that two plants are clearly over-utilizing labor relative to capital and two are under-utilizing. Only one plant is allocatively inefficient in both inputs relative to capital by an economically significant margin, plant 3, which is over-utilizing labor and under-utilizing energy relative to capital. The fact that its allocative inefficiencies are in opposite directions makes it likely that for plant 3 the allocation error is not in the amount of capital input relative to which we are measuring efficiency. Again, the figures make clear that some allocative inefficiency measures are distinguishable from others.

DOI: 10.1002/jae

Table III. Year 12 plant-specific allocative efficiency for energy

Plant	Post. mean	Post. median	Post. SE
1	1.160	1.160	0.044
2	1.176	1.170	0.037
3	0.978	0.990	0.069
4	1.003	1.000	0.024
5	0.821	0.820	0.023
6	1.030	1.030	0.027
7	1.001	1.000	0.025
8	0.947	0.950	0.025
9	0.907	0.910	0.021
10	0.875	0.870	0.026
11	0.928	0.930	0.034
12	0.984	0.980	0.027
13	0.952	0.950	0.026

Note: Measured relative to capital.

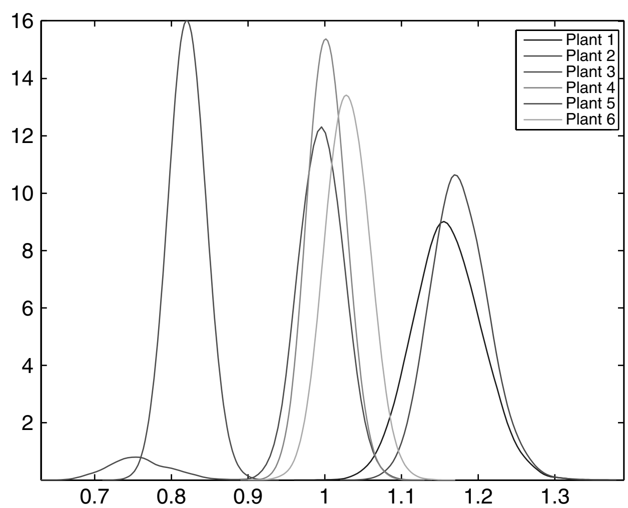

Figure 1. Plant-specific allocative inefficiency for energy

By contrast, estimation of the distance system with only industry-wide allocative efficiency parameters yields estimates very close to one for both labor and energy. These industry-wide estimates do not appear to suffer from aggregation bias (i.e., they are near the weighted average of the plant-specific values), but they clearly mask important inter-plant differences that are only observable when plant-specific allocative efficiency parameters are estimated.

Average technical efficiency scores, reported in Tables V and VI, show an enormous range of efficiencies across plants and over time. Table V shows the TE scores by plant for year 12 of our sample. Plant 1, a large run-of-river plant, is the most efficient for year 12, although that is the only year it is the most efficient. Six of the 13 plants have posterior median TE scores of 0.50 or lower, displaying considerable technical inefficiency. In fact, the posterior density region limits shown in Table V reveal that for three of the plants we can place a 95% probability level on their

DOI: 10.1002/jae

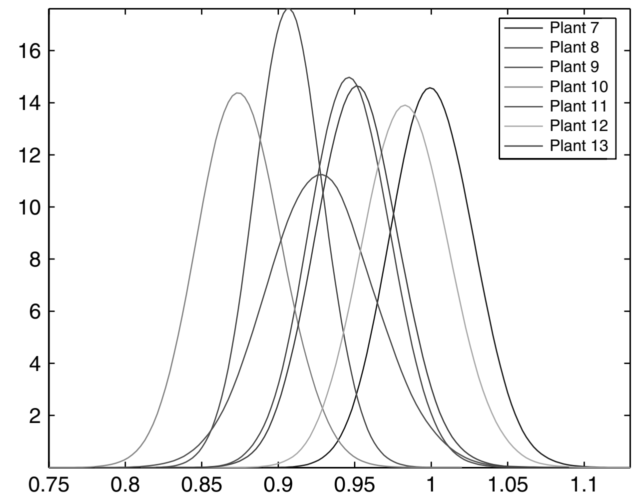

Figure 2. Plant-specific allocative inefficiency for energy

Table IV. Year 12 plant-specific allocative efficiency for labor

Plant	Post. mean	Post. median	Post. SE
1	0.936	0.940	0.046
2	0.726	0.730	0.019
3	0.981	0.980	0.026
4	0.983	0.980	0.024
5	1.038	1.040	0.035
6	1.004	1.000	0.030
7	1.034	1.030	0.027
8	1.112	1.110	0.027
9	1.024	1.020	0.026
10	0.980	0.980	0.030
11	1.141	1.140	0.045
12	0.767	0.770	0.021
13	0.965	0.960	0.027

Note: Measured relative to capital.

having a TE score of 0.50 or lower. The posterior density region limits also show that the TE scores are considerably more precise for some plants than for others. Table VI shows the changes in plant-specific TE scores over the sample period for three plants (since a table cannot easily display all 13 plant-specific measures for all 12 years). The results in Table VI clearly show how much the plant-specific TE scores are changing over time, in contrast to the k_{Et} and k_{Lt} values which were quite stable. Not only do plant-specific scores and rankings change rapidly, but the changes are not uniformly monotonic. Figure 5 displays the posterior distributions of TE scores for four representative plants, so the reader can see that some plants can be differentiated in terms of technical efficiency and that the distributions are, again, generally symmetric.

Table VII displays the year 9 estimates of the posterior medians for PC, TC, and EC for each plant. Year 9 was chosen as it is the first year that all plants are in the sample and to avoid

DOI: 10.1002/jae

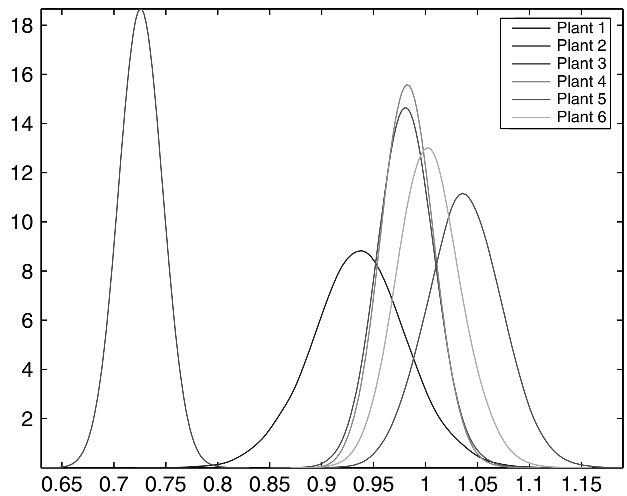

Figure 3. Plant-specific allocative inefficiency for labor

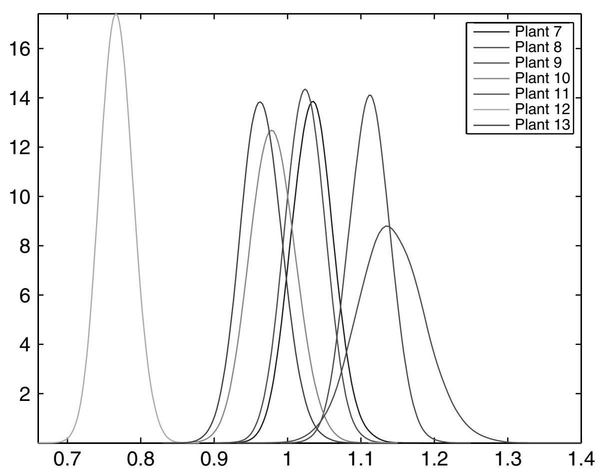

Figure 4. Plant-specific allocative inefficiency for labor

presenting only year 12 results. Table VIII contains the results of plant-specific PC for all years for three representative plants, to provide information on how these measures change over the sample period. Looking at the results in Table VII first, we see that PC varies widely from a high of 0.1254 to a low of 0.0195. Thus, between years 8 and 9, plant 2 increased overall productivity faster than plant 5 by a factor of six. TC scores are necessarily very similar across plants since the only difference is caused by the frontier moving outward at slightly different speeds along the separate rays from the origin that reflect each plant's specific input ratio. Median EC scores demonstrate that five plants are closing on the frontier, while eight plants are falling behind. Again, the variation is economically very significant, with plant 2 having an EC score of 0.0551 while plant 5's is -0.0487, almost equal in magnitude but opposite in sign.

DOI: 10.1002/jae

Table V. Year 12 plant-specific technical efficiencies

Plant	Lower limit	Median TE	Upper limit
1	0.98	1.00	1.00
2	0.72	0.82	0.93
3	0.48	0.56	0.63
4	0.39	0.45	0.53
5	0.35	0.42	0.50
6	0.47	0.55	0.63
7	0.59	0.70	0.82
8	0.37	0.43	0.49
9	0.41	0.48	0.54
10	0.33	0.39	0.44
11	0.40	0.49	0.59
12	0.61	0.79	0.99
13	0.50	0.64	0.76

Note: Lower and upper limits are for symmetric 90% posterior density regions.

Table VI. Selected plant-specific median technical efficiencies over time

Year	Plant 3	Plant 7	Plant 13
1	0.84	0.48	—
2	0.73	0.50	1.00
3	0.68	0.55	1.00
4	0.65	0.60	1.00
5	0.63	0.66	1.00
6	0.63	0.72	1.00
7	0.63	0.78	1.00
8	0.64	0.83	0.98
9	0.64	0.86	0.95
10	0.66	0.88	0.91
11	0.68	0.89	0.86
12	0.56	0.70	0.64

Now turning to Table VIII, which focuses on PC scores over time, one sees that, like TE, PC scores change rapidly across periods with plant-specific productivity growth being in spurts and stops—not a smooth process at all. As these three plants are all run-of-river, it is not too surprising that their efficiency can fluctuate from year to year. Figures 6, 7, and 8 show the posterior distributions for four plants of PC, TC, and EC, respectively, to provide additional insight into the empirical results.

The results show significant correlation between EC and TE scores. The computed correlation coefficient between EC and TE is 0.398. With 122 observations for annual EC among the 13 plants, that is very statistically significant. Thus, firms with higher TE scores also appear to be increasing their technical efficiency faster than the plants with lower current efficiency. This implies these plants are not converging, but instead the better-run plants are leaving the less efficient ones farther and farther behind.

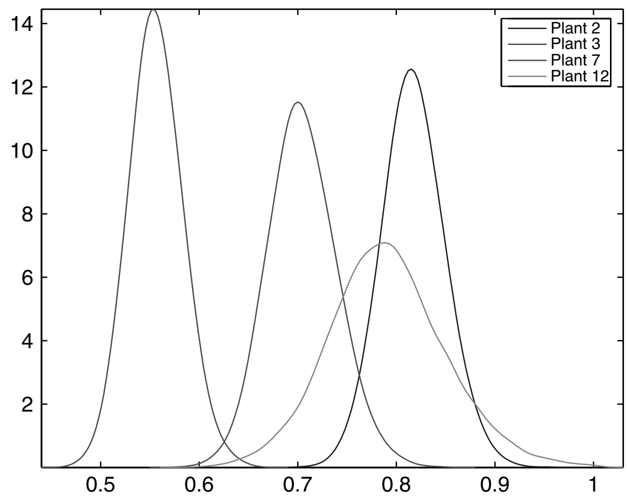

Figure 5. Plant-specific technical efficiency in year 12

Table VII. Year 9 plant-specific posterior medians for PC, TC, and EC

Plant	PC	TC	EC
1	0.0622	0.0673	−0.0045
2	0.1254	0.0714	0.0551
3	0.0779	0.0730	0.0053
4	0.0692	0.0809	−0.0113
5	0.0195	0.0692	−0.0487
6	0.0872	0.0706	0.0169
7	0.1003	0.0706	0.0301
8	0.0563	0.0787	−0.0219
9	0.0718	0.0770	−0.0048
10	0.0437	0.0730	−0.0285
11	0.0363	0.0699	−0.0320
12	0.0757	0.0746	0.0010
13	0.0342	0.0694	−0.0349

5. CONCLUSIONS

Random effects and non-Bayesian parametric approaches have dominated the literature on the estimation of allocative and technical efficiency. While the choice of functional form and the treatment of endogeneity are problematic with the random effects approach, we avoid specifying the distribution of the errors and can easily treat endogeneity with our parametric approach. Our Bayesian MCMC parametric method allows us to compute time-varying, plant-specific allocative inefficiency measures utilizing a limited information instrumental variable approach, which is analogous to Bayesian GMM with instruments. The allocative efficiency parameters are jointly estimated with the structural parameters of a translog distance function and its associated price equations. Our approach should allow for much richer investigation of the magnitude, precision, and distribution of allocative inefficiencies of plants or firms within an industry than could be

DOI: 10.1002/jae

Table VIII. Selected plant-specific median PC over time

Year	Plant 3	Plant 7	Plant 13
2	0.0362	0.1720	—
3	0.0335	.1312	0.0763
4	0.0381	0.1204	0.0591
5	0.0452	0.1127	0.0522
6	0.0455	0.1076	0.0553
7	0.0611	0.1090	0.0529
8	0.0708	0.1061	0.0424
9	0.0779	0.1003	0.0342
10	0.0837	0.0910	0.0151
11	0.0914	0.0826	0.0292
12	0.1789	0.1139	0.0840

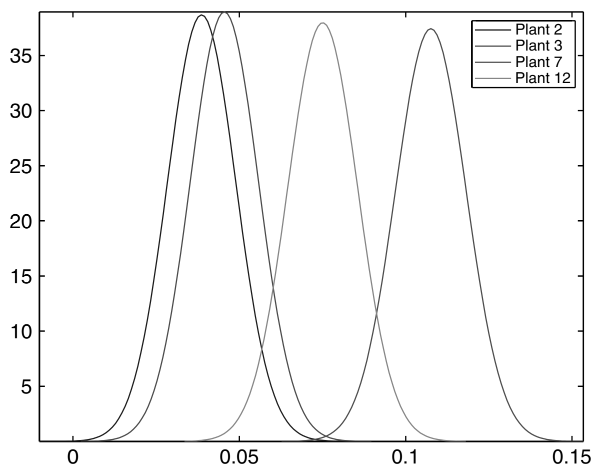

Figure 6. Plant-specific productivity changes for year 6

achieved in the previously used non-Bayesian parametric approaches, which typically were able to identify only industry-wide measures.

In our application, a panel of Chilean hydroelectric power plants do not appear to become more allocatively efficient over time by learning as they go. However, technical efficiency does stay constant or improve over time, as evidenced by zero or positive efficiency change for 12 of the 16 plants in our sample. The production frontier is being pushed outward at an estimated annual rate of 3% at the end of our sample. Most importantly, we found considerable differences in the location and precision of estimated allocative inefficiency measures across the plants in our sample, and found that industry-wide measures mask economically significant plant-specific heterogeneity.

ACKNOWLEDGEMENTS

We wish to thank Badi Baltagi, Chris Cornwell, John Geweke, Bill Horrace, Gary Koop, Subal Kumbhakar, Robin Sickles, Dek Terrell, Arnold Zellner, and seminar participants at Texas A&M

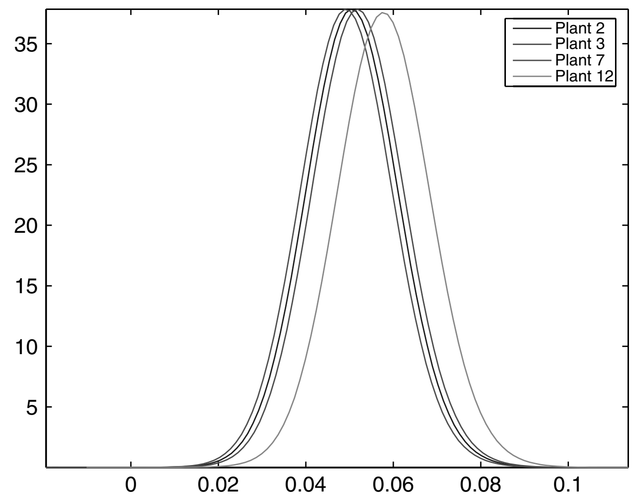

Figure 7. Plant-specific technical changes for year 6

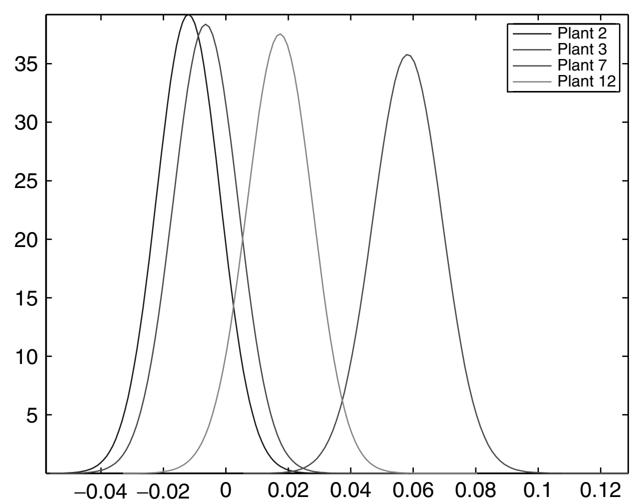

Figure 8. Plant-specific efficiency changes for year 6

University, Rice University, and the 2004 Southern Economic Association Meetings for helpful discussions and comments. We also thank Herman Van Dijk and three anonymous reviewers for helping to improve the paper through the editorial process.

REFERENCES

Atkinson SE, Cornwell C. 1994. Parametric measurement of technical and allocative inefficiency with panel data. *International Economic Review* **35**: 231–244.

Atkinson SE, Dorfman JH. 2005. Bayesian measurement of productivity and efficiency in the presence of undesirable outputs: crediting electric utilities for reducing air pollution. *Journal of Econometrics* **126**: 445–468.

Atkinson SE, Halabá C. 2005. Measuring productivity growth and efficiency of the Chilean electric utility industry. *Journal of Productivity Analysis* **23**: 245–273.

Atkinson SE, Halvorsen R. 1984. Parametric efficiency tests, economies of scale, and input demand in US electric power generation. *International Economic Review* **25**: 647–662.

Atkinson SE, Primont D. 2002. Measuring productivity growth, technical efficiency, allocative efficiency, and returns to scale using distance functions. *Journal of Econometrics* **108**: 203–225.

Atkinson SE, Färe R, Primont D. 2003. Stochastic measurement of firm inefficiency using distance functions. *Southern Economic Journal* **69**: 596–611.

Bauer PW. 1990. Recent developments in the econometric estimation of frontiers. *Journal of Econometrics* **46**: 39–56.

Casella G, George EI. 1992. Explaining the Gibbs sampler. *American Statistician* **46**: 167–174.

Cornwell C, Schmidt P, Sickles RC. 1990. Production frontiers with time series variation in efficiency levels. *Journal of Econometrics* **46**: 185–200.

Greene WH. 1980. On the estimation of a flexible frontier production model. *Journal of Econometrics* **13**: 101–115.

Green E, Strawderman W. 1996. A Bayesian growth and yield model for slash pine plantations. *Journal of Applied Statistics* **23**: 285–299.

Greene WH. 1997. Frontier production functions. In *Handbook of Applied Econometrics*. Vol. II: *Microeconomics*, Pesaran MH, Schmidt P (eds). Blackwell: Oxford; 81–166.

Griffin JE, Steel MFJ. 2004. Semiparametric Bayesian inference for stochastic frontier models. *Journal of Econometrics* **123**: 121–152.

Hansen B. 1982. Large sample properties of generalized method of moments estimators. *Econometrica* **50**: 1029–1054.

Haung CJ, Liu JT. 1994. Estimation of a non-neutral atochastic frontier production function. *Journal of Productivity Analysis* **5**: 171–180.

Kim JY. 2002. Limited information likelihood and Bayesian analysis. *Journal of Econometrics* **107**: 175–193.

Koop G, Poirier DJ. 2004. Bayesian variants of some classical semiparametric regression techniques. *Journal of Econometrics* **123**: 259–282.

Koop G, Tobias JL. 2006. Semiparametric Bayesian inference in smooth coefficient models. *Journal of Econometrics* **134**: 283–315.

Koop G, Osiewalski J, Steel FM. 1999. The components of output growth: a stochastic frontier analysis. *Oxford Bulletin of Economics and Statistics* **61**: 455–487.

Kumbhakar SC, Tsionas EG. 2005. The joint measurement of technical and allocative inefficiencies: an application of Bayesian inference in nonlinear random-effects models. *Journal of the American Statistical Association* **100**: 736–747.

LaFrance J. 1999. Inferring the nutrient content of food with prior information. *American Journal of Agricultural Economics* **81**: 728–734.

Lim H, Shumway CR. 1992. Profit maximization, returns to scale, and measurement error. *Review of Economics and Statistics* **74**: 430–438.

Newey WK, West KD. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* **55**: 703–706.

Tierney L. 1994. Markov chains for exploring posterior distributions (with discussion). *Annals of Statistics* **22**: 1701–1762.

Zellner A. 1998. The finite sample properties of simulataneous equations? Estimates and estimators: Bayesian and non-Bayesian approaches. *Journal of Econometrics* **83**: 185–212.

Zellner A, Highfield RA. 1988. Calculation of maximum entropy distributions and approximation of marginal posterior distributions. *Journal of Econometrics* **37**: 195–209.

Zellner A, Tobias J. 2001. Further results on Bayesian method of moments analysis of the multiple regression model. *International Economic Review* **42**: 121–140.

Zellner A, Bauwens L, Van Dijk HK. 1988. Bayesian specification analysis and estimation of simultaneous equation models using Monte Carlo methods. *Journal of Econometrics* **38**: 39–72.

DOI: 10.1002/jae