Joint Estimation of Production and Cost Models, With an Application to Electricity Distribution

Preliminary Draft, Not for Quotation*

Dimitrios Dimitropoulos† and Adonis Yatchew‡

Department of Economics, University of Toronto

August 11, 2014

Abstract

Although production and cost functions are mathematical duals – they both contain the same information about the underlying technology – practitioners typically estimate only one or the other. This paper proposes an approach for joint estimation of production and costs. Combining such quantity and price data has the effect of adding statistical information without introducing additional parameters into the model, thus allowing one to obtain greater precision of estimation. Our estimator optimizes a GMM-type objective function and simultaneously produces internally consistent parameter estimates for both the production function and the cost function. We show how to account for the presence of certain types of simultaneity and measurement error, and we extend the framework to a multi-output setting. The methodology is applied to data on 73 Ontario distributors for the period 2002-2012. Our models incorporate business condition variables such as customer density. We estimate productivity growth to be approximately *negative* 1% per year over the period. As expected, the joint model results in a substantial improvement in the precision of parameter estimates.

---

* Comments and suggestions welcome.

† Contact Address: Department of Economics, University of Toronto. Office No 320. 150 St. George Street, Toronto, Ontario, M5S 3G7, Canada. E-mail: d.dimitropoulos@utoronto.ca

‡ Contact Address: Department of Economics, University of Toronto. Office No 278. 150 St. George Street, Toronto, Ontario, M5S 3G7, Canada. E-mail: yatchew@chass.utoronto.ca

1 Introduction

From the economist's point of view, production functions and cost functions are duals -- knowledge of one permits recovery of the other. From the perspective of the empirical researcher, both quantity and price information should be useful in obtaining estimates of the underlying technology and cost structures. Standard cost models often incorporate factor share data to improve parameter estimates. Production function models may incorporate factor price data to constrain the slopes of isoquants. Both of these sources of derivative data can lead to substantial improvements in efficiency; in nonparametric settings, they can even increase the rates of convergence of estimators.

Nevertheless, practitioners overwhelmingly estimate either a production function or a cost function. The main reason is that, except in the simplest specifications such as the Cobb-Douglas class, there is an absence of tractable models for both production and costs. For example, a translog model of production does not have an analytical solution for its cost counterpart. Furthermore, there may be data availability, simultaneity and error in variables issues.

We seek to challenge these barriers. We take the production function as our primitive, and implement numerical optimization techniques to solve for the associated cost function and its derivatives i.e., the factor cost-shares. Combining the production function together with the cost function and its derivatives has the effect of adding informative data without introducing additional parameters into the model, thus allowing one to obtain greater precision of estimation. Our estimator optimizes a GMM-type objective function and produces internally consistent parameter estimates for both the cost function and the production function. Moreover, the framework permits tests of consistency of data on the cost function with data on the production function. We also show how to account for the presence of certain types of simultaneity and measurement error within the joint estimation framework.

We are motivated by the use of cost function data in regulating the distribution segment of the electricity industry. Since distributor cost functions typically involve more than one dimension of output, we extend the framework to a multi-output setting and apply it to data on 73 Ontario distributors for the period 2002-2012. In addition to the usual input price and output quantity data, our models incorporate variables that reflect customer density, and technological growth.

Our baseline model is a Cobb-Douglas specification for the primal, which then yields a Cobb-Douglas specification for the dual. We find that primal and dual estimates are qualitatively consistent with each other; of the two output variables – number of customers and deliveries per customer – the former has a much larger impact on production decisions than the latter; customer density has a material impact on production. In these particular data, productivity growth is found to be *negative* at a rate of approximately 1% per year over the 2002-2012 period. Moreover, correcting for simultaneity and measurement error has a significant impact on parameter estimates. We then extend estimation to a translog specification for the production function and numerically solve for estimates of the cost function and its derivatives.

The remainder of the paper is organized as follows. Section 2 reviews the literature on primal and dual estimation and places our work within this context. Section 3 introduces the model and empirical implementation. In Section 4, we begin with a brief introduction to the data and provide descriptive statistics. We then present and compare the results from both separate and joint estimation of the primal and dual for the case of the Cobb-Douglas production technology. Section 4.3 extends estimation to a more general framework and presents the results for the case where the primal consists of a translog specification. Section 5 concludes.

2 Literature Review

There has been much debate about which methodology, the primal or the dual, is best for the study of the productivity of modern industrial organizations (see, for example, Mundlak (1996)). Considerable attention has been devoted in the literature to the estimation of production functions, and the endogeneity issue first described by Marshack and Andrews (1944), stemming from idiosyncratic productivity shocks. Mundlak (1961) and Mundlak and Hoch (1965) were the first to propose a solution to the problem by using the standard ‘within’ panel data estimator in the context of agriculture production functions. Nevertheless, this approach has significant limitations, particularly when applied to other industries, such as manufacturing, where the assumption of a time-invariant productivity shock is unrealistic. Moreover, the within estimator is known to suffer from attenuation bias when the signal to noise ratio in the data is low.

Since then, two different approaches have been taken to the estimation of production function in the presence of simultaneity. One approach is to use dynamic panel data models that use lagged values of inputs as instruments for current values. Based on assumptions on adjustment costs and the timing of input decisions, Blundell and Bond (2000) extend the Arellano and Bond (1991) dynamic panel data estimator and derive two sets of moment conditions for the production function equation: lagged first differences of inputs are valid instruments in the equation in the "levels equation", and lagged levels of inputs are valid in instruments in the "first differenced" equation. In practice, the Blundell-Bond estimator has performed very well. In particular, it exploits cross-sectional variability, thus mitigating the problem of amplified measurement error in the time differences.

A second approach is to make assumptions on the covariance structure of the productivity shock, together with assumptions about the timing of input decisions, and derive a control function that can used to condition out serially correlated productivity shocks. Based on assumptions about the time-to-build nature of capital, Olley and Pakes (1994) use the investment policy function from a firm's dynamic programming problem to control for correlation between inputs and an idiosyncratic productivity shock.

Levinsohn and Petrin (2003), motivated by concerns about Olley and Pakes' assumption that investment is strictly monotonic in the productivity shock, make use of a non-dynamic intermediate ("materials") input demand equation proxy, which Levinsohn and Petrin argue is more likely to be monotone.

Levinsohn and Pakes apply their estimator to study the productivity of Chilean manufacturers across an array of industries.

Ackerberg, Caves and Frazer (2006) critique the Levinsohn and Petrin (and to a lesser extent Olley and Pakes) approach over concerns that, if their assumptions did hold true, then there should be exact collinearity in the first stage of the estimator. Thus, Levinsohn and Petrin should not be able to identify the parameters of their model. Ackerberg et al. go on to provide alternative estimator, which abandons the two-step approach to estimation taken by previous applications of the control function framework, and instead estimates all the parameters of the production function in a single stage.

On the estimation of cost functions, a seminal paper in the econometric analysis of the cost structure of the electricity industry is Christensen and Greene (1976). They use Zellner's (1962) Seemingly Unrelated Regressions model, along with a transcendental logarithmic specification for the cost function, to study economies of scale in US electric power generation. There is also a growing literature on the cost structure of electricity transmission and distribution. Early studies on US utilities by Meyer (1975), Neuberg (1977), and Huettner and Landon (1978), and considerably more sophisticated studies by Henderson (1985), Roberts (1986), and Nelson and Primeaux (1988) all report evidence of scale economies for US electric utilities.

More recent studies, although differing in methodology, data availability, and choice of variables, find evidence of distribution segment economies in a wide range of jurisdictions. These include: Hjalmarsson and Veiderpass (1992) on Sweden, Giles and Wyatt (1993) on New Zealand, Burns and Weyman-Jones (1996) on England and Wales, Filippini (1996, 1998) and Farsi and Filippini (2004) on Switzerland, Salvanes and Tjotta (1998) on Norway, Yatchew (2000) on Ontario, Canada, Kwoka (2005) on the US, and Farsi et al. (2010) on France.

A separate, though related literature, uses cost function estimation to calibrate a productivity index, in order to gauge the productivity trends in electricity distribution. Scully (1999) estimates the cost structure of electricity distributors in New Zealand, and then uses a Tornqvist index to gauge total factor productivity of New Zealand electricity distributors during the 1982 to 1994 period. Similar approaches have been taken by See and Coelli (2009) on Malaysia between 1975 and 2005, Goto and Sueyoshi (2009) on Japan between 1983 and 2003, Arcos and de Toledo (2009) on Spain between 1987 and 1997, and Dimitropoulos and Yatchew (2014) on Ontario between 2002 and 2012.

Several authors have made attempts to reconcile the two methodologies, and thus jointly estimate the primal and the dual, including e.g., Coen and Hickman (1970), McIntosh and Sims (1997), Paris and Caputo (2004), and Clark, Cechura and Thibodeau (2013). However, heretofore these studies have either not taken as serious the simultaneity issue, which has been made front and center by primal estimation, or have used various other constructs for the production set (e.g. distance functions) and not the production function proper. However, when data allow for the estimation of both these functions, then more efficient, and internally consistent, estimates can be obtained via joint estimation. Thus, our paper seeks to make the link between production function and cost function estimation as is proposed by duality theory.

The starting point of our analysis is closest to Coen and Hickman (1970), or Paris and Caputo (2004). However, we consider the case of a multi-output production technology. Coen and Hickman use factor demands to constrain the estimation of the production function. We estimate the production function, jointly with the cost function and the factor demands.

On the other hand, while the focus of Paris and Caputo is measurement error in output, we also take into consideration the simultaneity problem with inputs. Moreover, though both of us both of specify a error components for an arbitrary primal, our approach is to define a Generalized Methods of Moments estimator of the joint production and cost model, while Paris and Caputo use Nonlinear Seemingly Unrelated Regression.

3 The Model

We study the joint estimation of production and cost decisions of multi-product firms in the context of electricity distribution. While production functions and cost functions can each be independently estimated, when data allows for the estimation of both these functions, then more efficient, and internally consistent, estimates can be obtained via joint estimation. We take as our primitive the production function of a firm, and then derive the cost function implied by the firm's cost minimization problem. Combining production and cost data then allows us to derive a set of regression equations which can then be jointly estimated.

3.1 The Production Possibilities and Costs of an Electricity Distributor

Consider an industry consisting of J firms with identical production technologies. For a typical firm in the industry, the production set Y describes the set of technologically feasible net outputs of the firm's production process. If the production process is such that the set of outputs is distinct from the set of inputs, and transformation is separable in inputs and outputs, then the production set of the firm can be represented as

$$Y = \{(z,q) \in R^{L+M} : G(q) \leq F(z)\} \tag{1}$$

where q is a vector of the firm's M outputs, z is the vector of the firm's L inputs, the function $F{:}R^M \to R^P$ is a production function, and the function $G{:}R^L \to R^P$ is an output function. We assume that both the production function and the output function are single valued and, in particular, interpret $G{:}R^L \to R$ as an "output index" function.

Define the output index $Q \equiv G(q)$. In unregulated markets, the typical firm can select q, and thus Q. However, this is not so for an electricity distributor that is required to serve all customers in its service territory, and to deliver all of its customers' electricity demands. Thus, a typical electricity distributor regards $Q \equiv G(q)$ as fixed, and its only decision is how to choose its vector of inputs to meet this level of output. We assume that the electricity distributor takes as given the vector of input prices *w*, and the decision problem facing the electricity distributor is to choose its levels of inputs to minimize costs:

$$\min_{z \in R^N} w^{\intercal} z \quad s.t.\ Q \le F(z). \tag{2}$$

3.2 Joint Estimation of Production and Costs: A Baseline Model

We observe an industry consisting of *J* firms, over *T* years, indexed by *j* and *t* respectively. We assume that the output of each firm is two dimensional: number of customers *N*, and total deliveries of electricity *D*, in kilowatt hours (kWh). Output is produced using two inputs: labour *L* purchased in a competitive market at a wage rate *w*, and capital *K* obtained at a rental rate *r*. We assume a parametric production function $F(L,K;\alpha)$, which describes how these inputs are transformed into output, and a parametric output function $G(N,D;\theta)$, which aggregates outputs into an index of total output.[1] The production process for distributor *j* and year *t* can be described by the production constraint

$$G(N_{jt}, D_{jt};\theta) = F(L_{jt}, K_{jt};\alpha) + e_{jt} \tag{3}$$

where e_{jt} is a latent error. We assume that L_{jt} and K_{jt} are both non-dynamic, and that a firm observes both N_{jt} and D_{jt} at the time that it selects its input vector. However, e_{jt} is realized only after (L_{jt}, K_{jt}) has been selected. Here, the interpretation is that e_{jt} represents *ex post* efficiencies, or deficiencies, in production that occur at the time that L_{jt} and K_{jt} are actually applied to the production of Q_{jt}. *Ex ante*, L_{jt} and K_{jt} should be just sufficient to produce total output Q_{jt}. However, *ex post*, plans are imperfectly implemented, and thus either the firm may produce too little output, in which case the firm must acquire extra resources to meet its required production, or too much output, in which case the firm must dispose

[1] More generally, both the production and output function can be vector-valued to allow for the case of multiple production processes. We are motivated to make the stronger assumption of single-valuedness on the basis that, in our empirical application below, the regulator makes precisely this assumption for purposes of setting the regulatory rule.

of some resources.[2] We assume that these additional resources are purchased in competitive markets, and that an electricity distributor incurs a cost of f_{jt} to compensate for the discrepancy e_{jt}.

Let I_{jt} denote the information set of the firm when choosing L_{jt} and K_{jt}. Since the firm does not observe e_{jt}, and hence f_{jt}, at this time, the firm must form expectations about their values. We assume

$$E[e_{jt} \mid I_{jt}] = 0 \tag{4}$$

$$E[f_{it} \mid I_{jt}] = 0. \tag{5}$$

Given the required level of total output it must produce, the cost minimization problem facing the firm is[3]

$$\min_{L_{jt}, K_{jt}} w_{jt} L_{jt} + r_{jt} K_{jt} \quad s.t.\ G(N_{jt}, D_{jt}; \theta) \leq F(L_{jt}, K_{jt}; \alpha). \tag{6}$$

Let $L(w_{jt}, r_{jt}, N_{jt}, D_{jt}; \alpha, \theta)$ and $K(w_{jt}, r_{jt}, N_{jt}, D_{jt}; \alpha, \theta)$ be the firm's conditional factor demands that solve the problem in (6). Then, the *ex post* cost function of the firm is

$$C(w_{jt}, r_{jt}, N_{jt}, D_{jt}; \alpha, \theta) = w_{jt} L(w_{jt}, r_{jt}, N_{jt}, D_{jt}; \alpha, \theta) + r_{jt} K(w_{jt}, r_{jt}, N_{jt}, D_{jt}; \alpha, \theta) + f_{jt}\ . \tag{7}$$

Moreover, defining SL_{jt} and SK_{jt} as the cost shares[4] of L_{jt} and K_{jt}, respectively, by Shephard's lemma

$$SL_{jt} = \frac{\partial}{\partial w_{jt}} \ln C(w_{jt}, r_{jt}, N_{jt}, D_{jt}; \alpha, \theta) \tag{8}$$

$$SK_{jt} = \frac{\partial}{\partial r_{jt}} \ln C(w_{jt}, r_{jt}, N_{jt}, D_{jt}; \alpha, \theta). \tag{9}$$

Under the zero conditional mean assumption (4), the parameters (α, θ) are identified in equation (3) from the production constraint alone. Likewise, under the zero conditional mean assumption in (5), the

---

[2] For example, managers may have to step in to make up for any inefficiencies; or can go home early if there are efficiencies.

[3] The implicit assumption here is that, since both *L* and *K* are non-dynamic inputs, in each period *t* the firm makes its input decision myopically, minimizing the costs of production in period *t* alone. However, the general formulation of our model, as well as the estimation algorithm described in Section 3.2.2 is sufficiently flexible to apply to the case of dynamic decision making under uncertainty. We leave the estimation of the dynamic case for future research.

[4] That is, $SL_{jt} = w_{jt} L_{jt} / C_{jt}$ and $SK_{jt} = r_{jt} K_{jt} / C_{jt}$.

parameters (α, θ) are also identified by the cost model in (7) through (9). However, in a finite sample, each model will provide us with a different set of point estimates. Moreover, given that the error terms in (3) and (7) through (9) are likely to be correlated, we can gain a more efficient estimator by estimating the two models jointly, as was shown by Zellner (1962). In particular, combining the two models has the effect of adding statistical information without introducing additional parameters into the model.

3.2.1 The Cobb Douglas Case

Suppose that both the production and the output transformation function take a Cobb-Douglas form

$$\ln F(L_{jt}, K_{jt}; \alpha) = \alpha_0 + \alpha_1 \ln L_{jt} + \alpha_2 K_{jt} \tag{10}$$

$$\ln G(N_{jt}, D_{jt}; \theta) = \theta_1 \ln N_{jt} + \theta_2 \ln D_{jt} \tag{11}$$

then, the production constraint can be written as

$$\theta_1 \ln N_{jt} + \theta_2 \ln D_{jt} = \alpha_0 + \alpha_1 \ln L_{jt} + \alpha_2 \ln K_{jt} + e_{jt}. \tag{12}$$

Given data on $\{N_{jt}, D_{jt}, L_{jt}, K_{jt}\}$, by imposing distributional assumptions on (12) it is possible to obtain a consistent estimator of (α, θ) using, e.g., Full Information Maximum Likelihood.

On the other hand, the cost minimization problem

$$\min_{L_{jt}, K_{jt}} w_{jt} L_{jt} + r_{jt} K_{jt} \qquad s.t.\ \theta_1 \ln N_{jt} + \theta_2 \ln D_{jt} = \alpha_0 + \alpha_1 \ln L_{jt} + \alpha_2 \ln K_{jt} \tag{13}$$

implies an *ex post* cost function of the form

$$\ln C_{jt} = \beta_0 + \beta_1 \ln w_{jt} + \beta_2 \ln r_{jt} + \beta_3 \ln N_{jt} + \beta_4 \ln D_{jt} + f_{jt} \tag{14}$$

where $\beta_0 \equiv \ln[(\alpha_1/\alpha_2)^{\alpha_2/(\alpha_1+\alpha_2)} + (\alpha_2/\alpha_2)^{\alpha_1/(\alpha_1+\alpha_2)}] - \alpha_0/(\alpha_1+\alpha_2)$, $\beta_1 \equiv \alpha_1/(\alpha_1+\alpha_2)$, $\beta_2 \equiv \alpha_2/(\alpha_1+\alpha_2)$, $\beta_3 \equiv \theta_1/(\alpha_1+\alpha_2)$, and $\beta_4 \equiv \theta_2/(\alpha_1+\alpha_2)$. The capital share equation

$$SK_{jt} = \beta_2 + g_{jt} \tag{15}$$

where $g_{jt}^{(K)}$ is a latent error.[5] Again, given data on $\{C_{jt}, SK_{jt}, w_{jt}, r_{jt}, N_{jt}, D_{jt}\}$, the zero-conditional mean assumption in (5), together with assumptions on the covariance structure of the residuals in (14) and (15), implies that it is straight-forward to obtain a consistent estimator of (α, θ).

However, given the shared parameters across the production and cost models above, as well as potential correlations among the error terms, further efficiency gains may be available by estimating the model

$$e_{jt} = \theta_1 \ln N_{jt} + \theta_2 \ln D_{jt} - \alpha_0 - \alpha_1 \ln L_{jt} - \alpha_2 \ln K_{jt}$$

$$f_{jt} = \ln C_{jt} - \beta_0 - \beta_1 \ln w_{jt} - \beta_2 \ln r_{jt} - \beta_3 \ln N_{jt} - \beta_4 \ln D_{jt} \tag{16}$$

$$g_{jt} = SK_{jt} - \beta_2 \ .$$

To explicitly take into account the relationship between the error terms across equations we impose assumptions on the covariance structure of (e_{jt}, f_{jt}, g_{jt}). We specify a random effects model, where we allow for nonzero correlations within a particular firm, but impose zero correlations between firms. Fix distributor *j*, and consider the structure of second order moments of the disturbances. We assume that each error term in (16) follows an autoregressive model of order one

$$e_{jt} = \rho e_{jt-1} + u_{jt} \qquad |\rho| < 1 \tag{17}$$

$$f_{jt} = \phi f_{jt-1} + v_{jt} \qquad |\phi| < 1 \tag{18}$$

$$g_{jt} = \psi g_{jt-1} + w_{jt} \qquad |\psi| < 1 \tag{19}$$

where each of u_{jt}, v_{jt}, and w_{jt} are independent and identically distributed *over time*, with zero means, and variances σ_u^2, σ_v^2, and σ_w^2 respectively. However, for distributor *j*, there is contemporaneous nonzero correlation between u_{jt}, v_{jt}, and w_{jt}:

$$\operatorname{cov}\left(u_{jt}, v_{jt}\right) = \sigma_{uv} \text{ and } \operatorname{cov}\left(u_{jt}, v_{js}\right) = 0 \text{ for all } s \neq t \tag{20}$$

[5] Without loss of generality, the share equation for labour has been omitted since, as cost-shares must add up to unity, it is redundant.

$$\operatorname{cov}\left(u_{jt}, w_{jt}\right) = \sigma_{uw} \quad \text{and} \quad \operatorname{cov}\left(u_{jt}, w_{js}\right) = 0 \text{ for all } s \neq t \tag{21}$$

$$\operatorname{cov}\left(v_{jt}, w_{jt}\right) = \sigma_{vw} \quad \text{and} \quad \operatorname{cov}\left(u_{jt}, v_{js}\right) = 0 \text{ for all } s \neq t \;. \tag{22}$$

Together, these assumptions imply that for distributor j, e_{jt}, f_{jt}, and g_{jt} have means of zero, and variances of $\sigma_e^2 = \sigma_u^2 / \left(1-\rho^2\right)$, $\sigma_f^2 = \sigma_v^2 / \left(1-\phi^2\right)$, and $\sigma_g^2 = \sigma_w^2 / \left(1-\psi^2\right)$, respectively. The independence of firms yields a block diagonal structure for the covariance structure across distributors:

$$\Omega = \begin{bmatrix} \Omega_1 & 0 & \dots & 0 \\ 0 & \Omega_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \Omega_N \end{bmatrix} \tag{23}$$

with

$$\Omega_j = \left[\begin{array}{cccccccccccc}
\sigma_e^2 & \rho\sigma_e^2 & \dots & \rho^{T-1}\sigma_e^2 & \sigma_{ef} & \rho\sigma_{ef} & \dots & \rho^{T-1}\sigma_{ef} & \sigma_{eg} & \rho\sigma_{eg} & \dots & \rho^{T-1}\sigma_{eg} \\
\rho\sigma_e^2 & \sigma_e^2 & & \rho^{T-2}\sigma_e^2 & \phi\sigma_{ef} & \sigma_{ef} & & \rho^{T-2}\sigma_{ef} & \psi\sigma_{fg} & \sigma_{fg} & & \rho^{T-2}\sigma_{fg} \\
\vdots & & \ddots & & \vdots & & \ddots & & \vdots & & \ddots & \\
\rho^{T-1}\sigma_e^2 & \rho^{T-2}\sigma_e^2 & & \sigma_e^2 & \phi^{T-1}\sigma_{ef} & \phi^{T-2}\sigma_{ef} & & \sigma_{ef} & \psi^{T-1}\sigma_{fg} & \psi^{T-2}\sigma_{fg} & & \sigma_{fg} \\
\\
\sigma_{ef} & \phi\sigma_{ef} & \dots & \phi^{T-1}\sigma_e^2 & \sigma_f^2 & \phi\sigma_f^2 & \dots & \phi^{T-1}\sigma_f^2 & \sigma_{fg} & \phi\sigma_{fg} & \dots & \phi^{T-1}\sigma_{fg} \\
\rho\sigma_{ef} & \sigma_{ef} & & \phi^{T-2}\sigma_e^2 & \phi\sigma_f^2 & \sigma_f^2 & & \phi^{T-2}\sigma_f^2 & \psi\sigma_{fg} & \sigma_{fg} & & \phi^{T-1}\sigma_{fg} \\
\vdots & & \ddots & & \vdots & & \ddots & & \vdots & & \ddots & \\
\rho^{T-1}\sigma_{ef} & \rho^{T-2}\sigma_{ef} & & \sigma_{ef} & \phi^{T-1}\sigma_f^2 & \phi^{T-2}\sigma_f^2 & & \sigma_f^2 & \psi^{T-1}\sigma_{fg} & \psi^{T-2}\sigma_{fg} & & \sigma_{fg} \\
\\
\sigma_{eg} & \psi\sigma_{eg} & \dots & \psi^{T-1}\sigma_{eg} & \sigma_{fg} & \psi\sigma_{fg} & \dots & \psi^{T-1}\sigma_{fg} & \sigma_g^2 & \psi\sigma_g^2 & \dots & \psi^{T-1}\sigma_g^2 \\
\rho\sigma_{eg} & \sigma_{eg} & & \psi^{T-2}\sigma_{eg} & \phi\sigma_{fg} & \sigma_{fg} & & \psi^{T-2}\sigma_{fg} & \psi\sigma_g^2 & \sigma_g^2 & & \psi^{T-2}\sigma_g^2 \\
\vdots & & \ddots & & \vdots & & \ddots & & \vdots & & \ddots & \\
\rho^{T-1}\sigma_{eg} & \rho^{T-2}\sigma_{eg} & & \sigma_{eg} & \phi^{T-1}\sigma_{fg} & \phi^{T-2}\sigma_{fg} & & \sigma_{fg} & \psi^{T-1}\sigma_g^2 & \psi^{T-2}\sigma_g^2 & & \sigma_g^2
\end{array}\right] \tag{24}$$

and where $\sigma_{ef} = \sigma_{uv} / \left(1-\rho\phi\right)$, $\sigma_{eg} = \sigma_{ew} / \left(1-\rho\psi\right)$ and $\sigma_{fg} = \sigma_{vw} / \left(1-\phi\psi\right)$.

Estimation of the joint model (16) with the above covariance structure is performed using Hansen's (1982) Generalized Method of Moments. Define the vectors $x_{jt}^{(PF)} = (1, L_{jt}, K_{jt})^{\intercal}$ and $x_{jt}^{(CF)} = (1, w_{jt}, r_{jt}, N_{jt}, D_{jt})^{\intercal}$. Then, the zero conditional mean assumption from (4) implies the moment conditions

$$E\left[x_{jt}^{(PF)}e_{jt}\right]=0 \tag{25}$$

and the zero conditional mean assumption for (5) implies the moment conditions

$$E\left[x_{jt}^{(CF)}f_{jt}\right]=0$$
$$E\left[x_{jt}^{(CF)}g_{jt}\right]=0. \tag{26}$$

Our model has 5 parameters $(\alpha,\theta)=(\alpha_0,\alpha_1,\alpha_2,\theta_1,\theta_2)$, and thus is over-identified given the 13 moment conditions in (25) and (26). Our data are $y_{jt}=(N_{jt},D_{jt},C_{jt},SK_{jt},L_{jt},K_{jt},w_{jt},r_{jt})$ for *j=1,2,...,J* and *t=1,2,...T.* Define the function $g(y_{jt},\alpha,\theta)$ as the stacked moment conditions specified in (25) and (26). The GMM estimator of our model is the value $(\hat{\alpha},\hat{\theta})$ that solves

$$\min_{\alpha,\theta}\left(\frac{1}{J}\frac{1}{T}\sum_{j,t}g(y_{jt},\alpha,\theta)\right)'Q^{-1}\left(\frac{1}{J}\frac{1}{T}\sum_{j,t}g(y_{jt},\alpha,\theta)\right) \tag{27}$$

where $Q=E[g(y_{jt},\alpha,\theta)g(y_{jt},\alpha,\theta)^{\intercal}]$.

3.2.2 The General Case

For arbitrary production and output transformation functions, we cannot guarantee that the cost function has a closed form. Estimation of the parameters is nevertheless straightforward.

Algorithm *Joint Estimation of the Production and Cost Functions*

1. Select an initial parameter estimate $\mu_0=(\alpha_0,\theta_0)$.
2. For each firm *j* at time *t:*
    a. Compute the value of the production equation $e_{jt}=\ln G(N_{jt},D_{jt};\theta_0)-\ln F(L_{jt},K_{jt};\alpha_0)$.
    b. Solve (numerically) the cost minimization problem: $\min_{L_{jt},K_{jt}} w_{jt}L_{jt}+r_{jt}K_{jt}+f_{jt}$ subject to $G(N_{jt},D_{jt};\theta_0)\le F(L_{jt},K_{jt};\alpha_0)$.

c. Compute the value of the cost equation $f_{jt} = \ln C_{jt} - \ln \mathrm{C}(w_{jt}, r_{jt} N_{jt}, D_{jt}; \alpha_0, \theta_0)$, and the share equation $g_{jt} = SK_{jt} - \partial \ln \mathrm{C}(w_{jt}, r_{jt} N_{jt}, D_{jt}; \alpha_0, \theta_0) / \partial r$.

3. Define $g(y_{jt}, \alpha, \theta) = \left(x_{jt}^{(PF)} e_{jt}, x_{jt}^{(CF)} f_{jt}, x_{jt}^{(CF)} g_{jt} \right)^{\intercal}$. Construct the sample analog of the moment conditions $m(\mu_0) = \frac{1}{J} \frac{1}{T} \sum_{j,t} g(y_{jt}, \alpha_0, \theta_0)$ and compute the value of the GMM objective function $m(\mu_0)^{\intercal} Q^{-1} m(\mu_0)$.

4. Update the parameter estimate to $\mu_1 = (\alpha_1, \theta_1)$ using, e.g. a Newton-Raphson iteration: $(\alpha_1, \theta_1) = (\alpha_0, \theta_0) - \left(\frac{\partial^2 m(\mu_0)}{\partial \mu \partial \mu^{\intercal}} Q^{-1} m(\mu_0) + \frac{\partial m(\mu_0)}{\partial \mu}^{\intercal} Q^{-1} \frac{\partial m(\mu_0)}{\partial \mu^{\intercal}} \right)^{-1} \frac{\partial m(\mu_0)}{\partial \mu}^{\intercal} Q^{-1} m(\mu_0)$.

5. Iterate until the change in the parameter vector is below a pre-specified tolerance level.

3.3 Endogeneity Problems: Simultaneity and Measurement Error

The literature on production function estimation has elucidated the identification problems resulting from several sources of endogeneity. There is serious potential for simultaneity problems: observed inputs may be correlated with unobserved inputs or productivity shocks, and this correlation will introduce bias in our estimator. Likewise, there is potential for measurement error: for example, we observe the value of output *ex post*, after the realization of latent errors, which is different from the expected value of output upon which firms base their production decisions. For ease of exposition, our discussion is framed in the context of the Cobb Douglas case, but it is straightforward to extend the ideas to the general case.

3.3.1 The Simultaneity Problem

Consider a production function of the form

$$\ln F(L_{jt}, K_{jt}; \alpha) = \alpha_0 + \alpha_1 \ln L_{jt} + \alpha_2 K_{jt} + \omega_{jt} + e_{jt} \tag{28}$$

where ω_{jt} is a firm-specific "productivity shock" representing a production factor that is observed by the electricity distributor, but is unobserved by the econometrician. Since ω_{jt} affects productivity, the electricity distributor will take it into account when making its choices for L_{jt} and K_{jt}. Thus, as shown by Marshack and Andrews (1944), both L_{jt} and K_{jt} will be correlated with the productivity shock: $E[L_{jt}\omega_j] \neq 0$ and $E[K_{jt}\omega_j] \neq 0$.

The cost minimization problem facing an electricity distributor is now

$$\min_{L_{jt},K_{jt}} w_{jt}L_{jt} + r_{jt}K_{jt} \quad s.t.\ \theta_1 \ln N_{jt} + \theta_2 \ln D_{jt} = \alpha_0 + \alpha_1 \ln L_{jt} + \alpha_2 \ln K_{jt} + \omega_{jt} \tag{29}$$

and the resulting cost and share functions take the form

$$\ln C_{jt} = \beta_0 + \beta_1 \ln w_{jt} + \beta_2 \ln r_{jt} + \beta_3 \ln N_{jt} + \beta_4 \ln D_{jt} + \omega^*{}_{jt} + f_{jt} \tag{30}$$

$$SK_{jt} = \beta_2 + g_{jt} \tag{31}$$

where the β's are defined as before, and $\omega^*_{jt} = \omega_j / (\alpha_1 + \alpha_2)$.

Now, however, the GMM estimator we described in Section 3.2.1 is biased. In particular, several of the moment conditions from the production equation are no longer valid: defining the unobservable $e^*_{jt} = \omega_{jt} + e_{jt}$, we have $E[L_{jt}e^*_{jt}] \neq 0$ and $E[K_{jt}e^*_{jt}] \neq 0$.[6]

In the context of the estimation of production functions, several solutions to this simultaneity problem have been proposed. Mundlak (1961) and Mundlak and Hoch (1965) take a panel data approach using a within estimator. Blundell and Bond (2000) propose a dynamic panel data estimator based on assumptions of adjustment costs, and the timing of firms' choice decisions. Using similar assumptions, Olley and Pakes (1996), Levinsohn and Petrin (2003), and Ackerberg et al. (2006) propose various control function approaches to the simultaneity problem.

For a variety of reasons, IV solutions have not been broadly used in practice in the context of production function estimation. First, instruments (e.g. input prices) are not always available. Even when they are, it

[6] Note, the introduction of a productivity shock in the production function does not alter the moment conditions we have specified for the cost function. In particular, both w_{jt} and r_{jt} are factor prices determined on competitive markets, and thus unaffected by the productivity shocks incurred by any particular firm. Likewise, N_{jt} and D_{jt} are demand side variables which, as required by the regulator, the firm has no choice but to satisfy.

is often the case that these instruments only exhibit time series variation, and such variation alone is not sufficient for identification if, for example, we introduce a trend term in the production function. Lastly, the use of IV relies strongly on the assumption that the productivity shock itself is ω_{jt} is exogenous, i.e., firms do not affect the realization of ω_{jt}, nor its evolution over time. Nevertheless, despite all these problems, IV has the desirable feature that it relies on fewer auxiliary assumptions than other structural approaches.

In the current context, IV is also the intuitive approach. First, instruments are available: in particular, data on input prices are already being used for the construction of the cost function. Moreover, these input prices are indeed exogenous to the model, as they are assumed to be determined in competitive factor markets.

Define the vector $z_{jt}^{(PF)} = (1, w_{jt}, r_{jt})^{\intercal}$. The assumption that input prices are exogenous implies the moment conditions

$$E\left[z_{jt}^{(PF)} e_{jt}^{*} \right] = 0 . \qquad (32)$$

Moreover, the possibility of a simultaneity problem does not affect identification in the context of the cost function. Recalling that $x_{jt}^{(CF)} = (1, w_{jt}, r_{jt}, N_{jt}, D_{jt})^{\intercal}$, equations (26) become

$$\begin{aligned} E\left[x_{jt}^{(CF)} f_{jt}^{*} \right] &= 0 \\ E\left[x_{jt}^{(CF)} g_{jt}^{*} \right] &= 0 \end{aligned} \qquad (33)$$

Equations (32) and (33) define 13 moment conditions, and thus our model with 5 parameters is again over-identified.

3.3.2 The Error in Variables Problem

Heretofore, we have been assuming that the level of output that the regulator imposes on each distributor is perfectly observable to a firm at the time of its choosing L_{jt} and K_{jt}. However, this may not be true in practice. In particular, the value of output that a firm ends up producing by the end of year *t* is likely not equal to the value that the firm expected to produce when it made its factor decisions at the beginning of year. In this section, we consider the case where neither N_{jt} nor D_{jt} are part of the information set I_{jt} available to the firm when solving its decision problem.

Since output requirements are not precisely known when it chooses its cost-minimizing levels of L_{jt} and K_{jt}, the firm must form expectations about them. Define $n_{jt} \equiv E[\ln N_{jt} \mid I_{jt}]$ and $d_{jt} = E[\ln \mathrm{D}_{jt} \mid \mathrm{I}_{jt}]$. *Ex post*, we have $\ln N_{jt} = n_{jt} + \eta_{jt}^{(n)}$ and $\ln D_{jt} = d_{jt} + \eta_{jt}^{(d)}$.

Removing N_{jt} and D_{jt} from the information set I_{jt} does not affect the form of the production constraint: *ex post*, the firm must produce the output required by the regulator. However, the *true* cost function now becomes

$$\ln C_{jt} = \beta_0 + \beta_1 \ln w_{jt} + \beta_2 \ln r_{jt} + \beta_3 n_{jt} + \beta_4 d_{jt} + f_{jt}^{*} . \qquad (34)$$

A traditional error-in-variables problem arises in (35) where we use the observed values $\ln N_{jt}$ and $\ln D_{jt}$ in place of n_{jt} and d_{jt}.

If available, the usual solution to this error-in-variables problem is the use of instrumental variables. An intuitive place to look for such instruments for N_{jt} and D_{jt} would be on the demand side of the industry. In this context, a promising variable is peak demand, which in some models is incorporated as an output variable, however its inclusion can lead to collinearity issues. We have chosen to exclude it and rely upon it as an instrument. It is correlated with number of customer and deliveries, but is uncorrelated with the measurement error of these two variables.

For each distributor, we observe two important peak demands: a summer peak PS_{jt}, and a winter peak PW_{jt}. Define the vector $z_{jt}^{(CF)} = (1, w_{jt}, r_{jt}, PS_{jt}, PW_{jt})^{\mathsf{T}}$. As before, the assumption that input prices are exogenous implies the moment conditions for the error term in the production constraint

$$E[z_{jt}^{(PF)} e_{jt}^{*}] = 0 \tag{35}$$

where earlier we defined $z_{jt}^{(PF)} = (1, w_{jt}, r_{jt})^{\intercal}$. Moreover, for the cost model, we now have

$$\begin{aligned} E[z_{jt}^{(CF)} f_{jt}] &= 0 \\ E[z_{jt}^{(CF)} g_{jt}] &= 0 \end{aligned} \tag{36}$$

Joint estimation of (35) and (36) by Generalized Method of Moments yields a consistent estimator of the parameters, now accounting for the presence of both simultaneity and measurement error.

4 Data and Empirical Results

4.1 The Data

The data for our study consist of a balanced panel of 73 Ontario electricity distributors, with yearly observations spanning the years 2002 to 2012. Data collection was commissioned by the Ontario Energy Board, with the support and participation of electricity distributors, for the purposes of the 2014 rate-setting agenda.[7] The extensive data include, for each distributor: the quantities of outputs, including the number of customers, and total electricity deliveries in kWh; demand side variables measuring summer peak demand, and winter peak demand; the yearly total costs, as measured by the sum of capital and operation, maintenance and administrative (OM&A) costs (including labor); the cost shares of capital and OM&A; indices for the quantities of capital and OM&A; price indices for capital and OM&A; as well as a host of "business condition" variables, to be used as controls. A descriptive summary of the variables included in this study is provided in Table 1.

[7] For more information of how the data were constructed see Kaufman et al (2013) and Dimitropoulos and Yatchew (2014).

Table 1: Descriptive Statistics for Econometric Variables

	Sample Mean	Standard Deviation	Minimum	Maximum
Number of Customers (N)	63,344	162,116	1,119	1,221,411
Deliveries (D) in kWh	1,629,428,323	4,101,543,652	19,660,484	26,372,168,650
Total Cost (TC)	$40,534,167	$137,574,487	$226,629	$1,261,193,244
Capital Cost-Share (SK)	53.73%	11.19%	13.15%	76.96%
Capital (K)	1,454,092	4,889,710	5,495	40,822,675
OM&A (L)	134,433	455,737	1,600	4,412,604
Capital Input Price (WK)	$17.39	$0.57	$16.74	$18.33
OM&A Input Price (WL)	$106.35	$11.23	$77.33	$132.92
Distribution Line Length (LL) in km	2,718	13,913	21	119,500
Summer Peak Demand (PDS) in kW	311,301	734,140	3,226	5,018,278
Winter Peak Demand (PDW) in kW	283,236	704,771	4,356	4,420,214

Number of Firms in Sample: 73 Electricity Distributors. Sampling Frequency: Annually during the period 2002- 2012. Number Observations: 803.

4.2 Estimation of the Cobb-Douglas Model

This section describes the empirical results for the Cobb-Douglas model. Our empirical specification for the primitives of the model, are

$$\ln F(\alpha) = \mathrm{x}_{jt}^{\mathrm{T}}\alpha_0 + \alpha_{0,1} + \alpha_1 \ln L_{jt} + \alpha_2 \ln K_{jt} + e_{jt}^{*} \tag{37}$$

$$\ln G(\theta) = \theta_1 \ln N_{jt} + \theta_2 \ln D_{jt}, \ \ \theta_1 + \theta_2 = 1 \tag{38}$$

where x_{jt} is a vector of "business conditions" to account for observed differences in characteristics across utilities as well as productivity trends. In particular, we include a distribution line-length variable to capture factors associated with the spatial distribution of consumers in the distributor's service territory, and a linear trend to capture the effects of Hicks-neutral technology change, and overall production pressures.

The assumption that the parameters of the output transformation function sum to unity is to ensure that, e.g. a doubling of both measures of output translates into a doubling of the output index. The accompanying cost function is

$$\ln C(\alpha,\theta) = \mathrm{x}'_{jt}\beta_0 + \beta_{0,1} + \beta_1 \ln w_{jt} + \beta_2 \ln r_{jt} + \beta_3 \ln N_{jt} + \beta_4 \ln D_{jt} + f_{jt}^{*} \tag{39}$$

where $\beta_0 \equiv \alpha_0 / (\alpha_1 + \alpha_2)$, $\beta_{0,1} \equiv \bar{\beta}_{0,1} + ln[(\alpha_1/\alpha_2)^{\alpha_2/(\alpha_1+\alpha_2)} + (\alpha_2/\alpha_2)^{\alpha_1/(\alpha_1+\alpha_2)}] - \alpha_0/(\alpha_1+\alpha_2)$, $\bar{\beta}_{0,1}$ is the unconditional mean of f_{jt}^{*}, $\beta_1 \equiv \alpha_1/(\alpha_1+\alpha_2)$, $\beta_2 \equiv \alpha_2/(\alpha_1+\alpha_2)$, $\beta_3 \equiv \theta_1/(\alpha_1+\alpha_2)$, and

$\beta_4 \equiv \theta_2/(\alpha_1+\alpha_2)$.

4.2.1 Separate Estimation of the Primal and Dual

Tables 2 and 3 present parameter estimates from separate estimation of the production model and the cost model for the Cobb-Douglas case. The upper half of Table 2 reports the parameter estimates from estimation of the primal alone, while the lower half of the table reports the implied dual estimates with standard errors computed using the delta method. Similarly, Table 3 reports the parameter estimates from estimation of the dual alone, while the lower half of the table reports the implied estimates. Note that, without imposing additional assumptions about the stochastic processes of outputs, estimation of the primal does not permit consistent estimates of the parameters θ for the output function G. As such, to allow for comparisons across these tables, in both cases we calibrated the output function using parameter estimates from Dimitropoulos and Yatchew (2014).[8]

The first column of Table 2 reports the OLS estimates of the primal (i.e., production) parameters, and the second column reports the IV estimates. In the presence of simultaneity, we expect that the coefficient on OM&A, *lnL* ,to be biased upwards under OLS estimation, and the coefficient on capital, *lnK*, to be biased downwards. Our results confirm this, indicating that simultaneity is an important concern in the estimation of the primal.

The first column of Table 3 reports the OLS estimates of the dual (i.e., cost) coefficients, and the second column reports the IV estimates. In the presence of measurement error, the coefficient on output, *lnQ*, suffers from conventional attenuation bias. Again, our results are consistent with this, indicating that errors-in-variables is likely an important concern in the estimation of the dual.

Comparing the results in Tables 2 and 3, we see that both the primal and the dual yield plausible coefficient estimates and confirm that electricity distribution is a relatively capital-intensive industry. However, the primal and dual estimates of output and cost elasticities are sufficiently different to suggest that either the two models need to be estimated jointly to allow data-based calibration of the output function, or the Cobb-Douglas specification may not be sufficiently flexible for these data.

4.2.2 Joint Estimation of the Primal and Dual

Table 4 presents the results from joint estimation of the production and cost model. Recall that we take as our primitive the production set of a firm, specifying parametric forms for the output transformation

[8] Specifically, θ_1 the weight associated with the number of customers was 0.8049, and θ_2 the weight associated with deliveries was 0.1951.

function and production function, and then derive (analytically) the cost function implied by the first cost-minimization problem. The first column of Table 4 reports the GMM estimates of the output, production, and cost functions for the Cobb-Douglas model, where we ignore the possibilities of simultaneity and measurement error. The second column reports estimates of these coefficients where we allow for the possibility of simultaneity bias in the production equation, instrumenting for inputs using factor prices. The third column reports estimates of these coefficient allowing for both simultaneity and measurement errors in output variables.

The estimates obtained from each of the models are all of the expected sign and statistically significant. As expected, the joint estimates all have standard errors that are substantially smaller than the estimates obtained from separate estimation of the primal and dual. Across all models, we find that: subscriber numbers are of first order importance in the determination of output, relative to kWh delivered; electricity distribution is relatively more capital intensive than labor intensive; and there are diseconomies of scale in electricity distribution. Nevertheless, there are substantial differences between these estimates. In particular, as we first correct for simultaneity, and then again for measurement error, the importance of subscribers declines, while the estimated capital intensity of production increases. A Hausman test rejects the null hypothesis of no simultaneity or measurement error. We focus our discussion on the estimates in the third column.

The estimates of the elasticity shares from the output transformation function imply that a 10% increase in subscriber numbers will result in a 8.1% increase in the output index, ceteris paribus, while a similar increase in deliveries will result in only a 1,9% increase in the output index.[9] Turning to the production function, our estimates imply that a 10% increase in capital will result approximately in a 5.2% increase in output, while a similar increase in OM&A will result approximately in a 4.6% increase. Together, these point estimates imply a 10% increase in both capital and OM&A will only increase output by 9.8%.

[9] Note that, we have restricted these parameters to add to unity, so that a 10% increase in both dimensions of output will necessarily lead to a 10% increase in the output index.

A test of the constant returns to scale is not rejected. These findings have significant implications for the cost function. The resulting capital-price elasticity of costs is 0.53, while that for OM&A is 0.46, i.e. a 10% input price increase will yield a 5.3% increase in total costs in the case of capital, and a 4.6% increase in the case of OM&A.

Likewise, the estimates of the parameters from the output and production function together also have important implications for the cost function. Our estimates imply a cost elasticity for subscribes of 0.82, and for deliveries of 0.20. By adding these together, we find that implied scale elasticity of costs is 1.02. That is, ceteris paribus, a 10% increase in subscribers will be associated with a 8.2% increase in costs, while a similar increase in deliveries will result in a 2.0% increase in total cost. A 10% increase in both subscribers and deliveries will result in a 10.2% increase in total costs, i.e., scant evidence of decreasing costs.

To control for "business conditions", we included the (log of) total circuit distribution line as a variable in the production function. Interpreting the estimated effect of this variable is more straightforward in the context of the cost function: we find that a 10% increase in total circuit distribution line raises distribution cost by 3.3%. Total circuit distribution line reflects the spatial distribution of consumers over a distributor's licensed service territory, and the distributor's associated need to construct delivery systems that transport electrons directly to the premises of end-users.

Our estimate of the trend coefficient implies that exogenous time-trending production pressures have, ceteris paribus, resulted in a 1.19% yearly decrease in production efficiency. Strictly speaking, while the time trend in our cost function is interpreted as capturing Hicks-neutral technical change, i.e. one that does not affect the balance of labor and capital in the production process, it may also be absorbing market changes.[10] Thus, our results imply that, over the sample period, the Ontario electricity distribution sector

[10] Between 2002 and 2012, there certainly was significant market changes in the Ontario electricity distribution sector (see, e.g., Dimitropoulos and Yatchew (2014)), but a full appreciation of these changes is outside the scope of this study.

has been experiencing cost pressures at the rate of 1.22% per year over the sample period, possibly in order to improve the quality of electricity supply, the implementation of demand management or FIT programs.

4.3 Estimation of More General Forms of Production Possibilities

The estimation algorithm described in Section 3.2.2 is sufficiently general to allow for the estimation of more general forms of production possibilities.[11] In this section we consider the case of a transcendental logarithmic (translog) specification for the production function. This form is attractive because it places no *a priori* restrictions on substitution patterns, and it encompasses more standard functional forms of the production function as special cases.

4.3.1 Empirical Specification

We maintain the functional form for the output function, as specified in equation (40), but the production function is now

$$\ln F(\alpha) = \mathrm{x}_{jt}^{\intercal}\alpha_0 + \alpha_{1,1}\ln L_{jt} + \alpha_{2,1}\ln K_{jt} + \alpha_{1,2}\ln L_{jt}^2 + \alpha_{2,2}\ln K_{jt}^2 + \alpha_3 \ln L_{jt}\ln K_{jt} + \omega_{jt} + e_{jt} \qquad (40)$$

The translog specification requires that the approximation of the underlying cost function to be made at a local point, which in our case is taken to be the vector of means of the variables. That is, right-hand-side variables are normalized by their mean value, therefore centering our cost function at a notional 'average firm'. This allows us to interpret the coefficients of the 'first order' terms of the input variables, as well as the coefficients of business conditions, as elasticities for this 'average firm'.

The translog specification does not allow for an analytical solution for the cost-function or the share equations. Nevertheless, so long as the production equation is quasi-concave, a solution to the cost-

[11] Indeed, the algorithm in Section 3.2.2 allows for more general forms of the transformation functions *G* and *F*, or the choice problem (e.g. one involving dynamic variable selection) facing the firm.

minimization problem exists, and can be obtained numerically. Let $C(\mathrm{x}_{jt}, w_{jt}, r_{jt}, N_{jt}, D_{jt}; \alpha, \theta)$ be the solution of the cost-minimization problem, and let $SK(\mathrm{x}_{jt}, w_{jt}, r_{jt}, N_{jt}, D_{jt}; \alpha, \theta)$ be the associated cost-share of capital. Then we can augment equation (42) with[12]

$$\begin{aligned} \ln C_{jt} &= \ln C(\mathrm{x}_{jt}, w_{jt}, r_{jt}, N_{jt}, D_{jt}; \alpha, \theta) + f_{jt}^{*} \\ \ln SK_{jt} &= \ln SK(\mathrm{x}_{jt}, w_{jt}, r_{jt}, N_{jt}, D_{jt}; \alpha, \theta) + g_{jt} \ . \end{aligned} \qquad (41)$$

4.3.2 Joint Estimation of the Primal and Dual

Table 5 presents the results of joint estimation of the production and cost models for the translog case. The first column reports the GMM estimates of the output and production parameters ignoring the possibility of simultaneity and measurement error. The second column reports the estimates where we allow for the possibility of simultaneity bias in the production equation, and the third reports the estimates from model where we account for both simultaneity and measurement error of outputs.

Qualitatively, the results are similar as before. A distributor's number of customers are a more important factor in its output decisions than are its kWh delivered. Further, for the notional 'average firm' electricity distribution is relatively more capital intensive than labor intensive, and there is no evidence of economies of scale in electricity distribution.

However, there are material quantitative differences between the Cobb-Douglas and the translog estimates. The output elasticity share of subscribers increases to 0.87 as compared to 0.78 earlier. Thus, it would appear that subscriber numbers are overwhelmingly the most critical cost driver. Furthermore, for the notional average firm, the factor share of capital is 0.56 as compared to 0.52 in the joint Cobb-Douglas specification, with a commensurate decline in the labour factor share. Finally, the introduction

[12] The assumption is that the solution to the cost minimization problem is additively separable in f_{jt} and ω_{jt}.

of higher-order and interaction terms allows for areas of increasing, constant, and decreasing returns to scale over the population of distributors which display considerable variation in size.

Focusing on Table 5, we again see that accounting for endogeneity has significant impact on our parameter estimates. For the output function, correcting simultaneity first increases the relative importance of electricity deliveries, but correcting for measurement error brings it back down slightly. For the production function, accounting for, first, simultaneity, and then measurement error, has a monotonic effect on the first-order parameters, each time increasing the relative intensity of capital. A Hausman test rejects the null hypothesis of no simultaneity or measurement error.

5 Conclusion

The economic theory of production and cost functions has evolved within a unified framework through the use of duality theory. Results such as the envelope theorem (of which Shephard's Lemma is a special case) provide an elegant link between quantities and prices.

In contrast the econometrics of production and cost functions has been implemented largely along separate paths. The absence of tractable, mutually consistent specifications for the primal and the dual has been an important stumbling block. Joint estimation has been further hindered by issues such as simultaneity and measurement error which can affect one or both sides of the model.

This paper propose a methodology for joint estimation by selecting a parametric specification for the production function and simulating the implied form of the cost function. The estimator employs a GMM-type objective function and produces internally consistent estimators of the primal and the dual. We also propose methods for dealing with certain types of simultaneity and measurement error.

Our application involves panel data on 73 electricity distributors with a two-output production function to which our methods can be readily adapted. We implement two specifications, the Cobb-Douglas which has an analytic representation for both production and costs, and the translog production function, for which the cost function requires numerical simulation. Our results indicate significant benefits from joint estimation as standard errors decline significantly relative to separate estimation of costs and production.

References

1. Ackerberg, D., Caves, K., & Fraser, G. (2004). Structural identification of production functions. Mimeo. UCLA.
2. Ackerberg, D., Lanier Benkard, C., Berry, S., & Pakes, A. (2007). Econometric tools for analyzing market outcomes. Handbook of Econometrics Volume 6, 4171-4276.
3. Aguirregabiria, V. (2009). Econometric Issues and Methods in the Estimation of Production Functions. Mimeo.
4. Arcos, A., & de Toledo, P. A. (2009). An analysis of the Spanish electrical utility industry: Economies of scale, technological progress and efficiency. Energy Economics, 31(3), 473-481.
5. Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. The Review of Economic Studies, 58(2), 277-297.
6. Blundell, R., & Bond, S. (2000). GMM estimation with persistent panel data: an application to production functions. Econometric Reviews, 19(3), 321-340.
7. Burns, P., & Weyman-Jones, T. G. (1996). Cost Functions and Cost Efficiency in Electricity Distribution: A Stochastic Frontier Approach. Bulletin of Economic Research, 48(1), 41-64.
8. Christensen, L. R., & Greene, W. H. (1976). Economies of scale in U. S. electric power generation. The Journal of Political Economy, 84(4), 655.
9. Clark, J. S., Cechura, L., & Thibodeau, D. R. (2013). Simultaneous estimation of cost and distance function share equations. Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie, 61(4), 559-581.
10. Coen, Robert M., and Bert G. Hickman (1970). Constrained joint estimation of factor demand and production functions. The Review of Economics and Statistics, 52(3), 287-300.
11. Dimitropoulos, D., & Yatchew, A. (2014). Is Productivity Growth in Electricity Distribution Negative? An Empirical Analysis Using Ontario Data. Mimeo.

12. Farsi, M., & Filippini, M. (2004). Regulation and measuring cost-efficiency with panel data models: Application to electricity distribution utilities. Review of Industrial Organization, 25(1), 1-19.

13. Farsi, M., Filippini, M., Plagnet, M. A., & Saplacan, R. (2010). The Economies of Scale in the French Power Distribution Utilities. Mimeo. Centre for Energy Policy and Economics, Swiss Federal Institutes of Technology

14. Filippini, M. (1996). Economies of scale and utilization in the Swiss electric power distribution industry. Applied Economics, 28(5), 543-550.

15. Filippini, M. (1998). Are municipal electricity distribution utilities natural monopolies? Annals of public and cooperative economics, 69(2), 157-174.

16. Giles, D. E. A., & Wyatt, N. S. (1993). Economies of Scale in New Zealand Electricity Distribution Industry. In Models, Methods and Application of Econometrics, ed. by P. C. B. Phillips. Cambridge, MA. Blackwell.

17. Goto, M., & Sueyoshi, T. (2009). Productivity growth and deregulation of Japanese electricity distribution. Energy Policy, 37(8), 3130-3138.

18. Hansen, L. P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. Econometrica, 50(4), 1029-54.

19. Henderson, J. S. (1985). Cost estimation for vertically integrated firms: The case of electricity. In Analyzing the Impact of Regulatory Change in Public Utilities, ed. M. A. Crew. Lexington, MA: Lexington Books.

20. Hjalmarsson, L., & Veiderpass, A. (1992). Productivity in Swedish electricity retail distribution. Scandinavian Journal of Economics, 94, Supplement, 193-205.

21. Huettner, D. A., & Landon, J. H. (1978). Electric utilities: scale economies and diseconomies. Southern Economic Journal, 44(4), 883-912.

22. Kaufmann, L., Hovde, D., Kalfayan, J., & Rebane, K. (2013). Empirical research in support of incentive rate setting in Ontario: Report to the Ontario Energy Board. Pacific Economics Group Research, LLC.

23. Kwoka Jr, J. E. (2005). Electric power distribution: economies of scale, mergers, and restructuring. Applied Economics, 37(20), 2373-2386.

24. Levinsohn, J., & Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. Review of Economic Studies, 70(2), 317-341.

25. Marschak, J., & Andrews, W. H. (1944). Random simultaneous equations and the theory of production. Econometrica, 12(3/4), 143-205.

26. Meyer, R. A. (1975). Publicly owned versus privately owned utilities: A policy choice. The Review of Economics and Statistics, 57(4), 391-399.

27. Mundlak, Y. (1961). Empirical production function free of management bias. Journal of Farm Economics, 43(1), 44-56.

28. Mundlak, Y. (1996). Production function estimation: reviving the primal. Econometrica, 64(2), 431-438.

29. Mundlak, Y., & Hoch, I. (1965). Consequences of alternative specifications in estimation of Cobb-Douglas production functions. Econometrica, 33(4), 814-828.

30. Nelson, R. A., & Primeaux, W. J. (1988). The effects of competition on transmission and distribution costs in the municipal electric industry. Land Economics, 64(4), 338-346.

31. Neuberg, L. G. (1977). Two issues in the municipal ownership of electric power distribution systems. The Bell Journal of Economics, 8(1), 303-323.

32. Olley, G.S., & Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. Econometrica, 64(6), 1263-1298.

33. Roberts, M. J. (1986). Economies of Density and Size in the Production and Delivery of Electric Power. Land Economics, 62(4), 378-387.

34. Salvanes, K. G., & Tjotta, S. (1998). A test for natural monopoly with application to Norwegian electricity distribution. Review of Industrial Organization, 13(6), 669-685.

35. Scully, G. W. (1999). Reform and efficiency gains in the New Zealand electrical supply industry. Journal of Productivity Analysis, 11(2), 133-147.

36. See, K. F., & Coelli, T. J. (2009). The Effects of Competition Policy on TFP Growth: Some Evidence from the Malaysian Electricity Supply Industry. Mimeo. University of Queensland, School of Economics.

37. Yatchew, A. (2000). Scale economies in electricity distribution: A semiparametric analysis. Journal of applied Econometrics, 15(2), 187-210.

38. Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. Journal of the American Statistical Association, 57(298), 348-368.

Table 2: Separate Estimation of the Production Model
Cobb-Douglas Specification

	No Endogeneity	Simultaneity
Production Function		
Constant	11.7570 (0.0295)**	12.1170 (0.0432)**
lnL	0.3135 (0.0163)**	0.2975 (0.0377)**
lnK	0.4333 (0.0144)**	0.6439 (0.0295)**
Line Length	-0.5227 (0.0318)**	-0.2917 (0.0634)**
Trend	-0.0088 (0.0008)**	-0.0093 (0.0075)**
Implied Cost Function		
Constant	15.744 (0.2614)**	12.8715 (0.1161)**
lnW_L	0.4198 (0.0188)**	0.3160 (0.0373)**
lnW_K	0.5802 (0.0188)**	0.6840 (0.0373)**
lnQ	1.3391 (0.0241)**	1.0623 (0.0931)**
Line Length	0.6999 (0.0442)**	0.3100 (0.0551)**
Trend	0.0118 (0.0010)**	0.0099 (0.0080)
Account For:		
Simultaneity	No	Yes
Error in Variables	No	No

Standard errors in parentheses. *,** indicate statistical significance at 10% and 5% levels, respectively. Number of Observations: 803 - 73 electricity distributors annually over the period 2002- 2012.

Table 3: Separate Estimation of the Cost Model
Cobb-Douglas Specification

	No Endogeneity	Error in Variables
Cost Function		
Constant	12.657 (0.0303)**	12.691 (0.0478)**
lnW_L	0.5386 (0.0126)**	0.5388 (0.0352)**
lnW_K	0.4614 (0.0126)**	0.4612 (0.0352)**
lnQ	0.9602 (0.0151)**	1.0232 (0.0250)**
Line Length	0.5588 (0.0313)**	0.3196 (0.0679)**
Trend	0.0160 (0.0007)**	0.0110 (0.0112)
Implied Production Function		
Constant	13.1810 (0.1884)**	12.404 (0.2765)**
lnL	0.4805 (0.0150)**	0.4507 (0.0357)**
lnK	0.5609 (0.0157)**	0.5266 (0.0362)**
Line Length	-0.3039 (0.0341)**	-0.2852 (0.0669)**
Trend	-0.0097 (0.0007)**	-0.0091 (0.0109)
Account For:		
Simultaneity	No	No
Error in Variables	No	Yes

Standard errors in parentheses. *,** indicate statistical significance at 10% and 5% levels, respectively. Number of Observations: 803 - 73 electricity distributors annually over the period 2002- 2012.

Table 4: Joint Estimation of the Production and Cost Models
Cobb-Douglas Specification

	No Endogeneity	**Simult.**	**Simult & EV**
		Production Model Estimates	
Output Function			
lnN	0.9659 (0.0146)**	0.8703 (0.0200)**	0.8054 (0.0214)**
lnD	0.0341 (0.0146)**	0.1297 (0.0200)**	0.1947 (0.0214)**
Production Function			
Constant	11.4800 (0.1460)**	12.5080 (0.2069)**	13.1620 (0.2201)**
lnL	0.4558 (0.0036)**	0.4545 (0.0093)**	0.4551 (0.0150)**
lnK	0.4846 (0.0043)**	0.5220 (0.0104)**	0.5230 (0.0163)**
Line Length	-0.3206 (0.0026)**	-0.3169 (0.0027)**	-0.3241 (0.0041)**
Trend	-0.0101 (0.0006)**	-0.0113 (0.0016)**	-0.0119 (0.0039)**
		Implied Cost Function Estimates	
Constant	12.7460 (0.0658)**	12.694 (0.0258)**	12.698 (0.0339)**
lnW_L	0.4847 (0.0033)**	0.4654 (0.0100)**	0.4653 (0.0160)**
lnW_K	0.5153 (0.0033)**	0.5346 (0.0100)**	0.5347 (0.0160)**
lnN	1.0271 (0.0165)**	0.8912 (0.0207)**	0.8234 (0.0215)**
lnD	0.0363 (0.0155)**	0.1328 (0.0205)**	0.1990 (0.0218)**
Line Length	0.3410 (0.0021)**	0.3245 (0.0028)**	0.3314 (0.0041)**
Trend	0.0108 (0.0006)**	0.0115 (0.0017)**	0.0122 (0.0040)**
Account For:			
Simultaneity	No	Yes	Yes
Error in Variables (EV)	No	No	Yes

Standard errors in parentheses. *,** indicate statistical significance at 10% and 5% levels, respectively. Number of Observations: 803 -- 73 electricity distributors annually over the period 2002- 2012.

Table 5: Joint Estimation of the Production and Cost Models
Translog Specification

	No Endogeneity	**Simult.**	**Simult & EV**
		Production Model Estimates	
Output Function			
lnN	0.8697 (0.0153)**	0.8425 (0.0337)**	0.8676 (0.0603)**
lnD	0.1303 (0.0153)**	0.1575 (0.0337)**	0.1324 (0.0603)**
Production Function			
Constant	12.496 (0.1542)**	12.839 (0.3445)**	12.571 (0.6122)**
lnL	0.4341 (0.0040)**	0.4152 (0.0042)**	0.4122 (0.0285)**
lnK	0.5430 (0.0036)**	0.5645 (0.0041)**	0.5624 (0.0264)**
lnL^2	0.0777 (0.0067)**	0.0301 (0.0073)**	0.0243 (0.0448)
lnK^2	0.1214 (0.0043)**	0.0975 (0.0044)**	0.0945 (0.0219)**
*lnL*lnK*	-0.1052 (0.0060)**	-0.0832 (0.0069)**	-0.0787 (0.0337)**
Line Length	-0.3209 (0.0052)**	-0.2691 (0.0056)**	-0.3138 (0.0097)**
Trend	-0.0138 (0.0004)**	-0.0092 (0.0005)**	-0.0104 (0.0015)**
Account For:			
Simultaneity	No	Yes	Yes
Error in Variables	No	No	Yes

Standard errors in parentheses. *,** indicate statistical significance at 10% and 5% levels, respectively. Number of Observations: 803 -- 73 electricity distributors annually over the period 2002- 2012.