Energy Economics 39 (2013) 233–238

Contents lists available at SciVerse ScienceDirect

Energy Economics

journal homepage: www.elsevier.com/locate/eneco

Energy substitution: When model selection depends on the focus

Peter Behl [a], Holger Dette [a], Manuel Frondel [a,b,*], Harald Tauchmann [b,c]

[a] Ruhr University Bochum (RUB), Germany
[b] Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI), Germany
[c] Universität Erlangen-Nürnberg, Germany

A R T I C L E I N F O

Article history:
Received 5 July 2012
Received in revised form 4 April 2013
Accepted 27 April 2013
Available online 17 May 2013

JEL classification:
C3
D2

Keywords:
Information criteria
Translog cost function
Cross-price elasticities

A B S T R A C T

In contrast to conventional model selection criteria, the Focused Information Criterion (FIC) allows for the purpose-specific choice of model specifications. This accommodates the idea that one kind of model might be highly appropriate for inferences on a particular focus parameter, but not for another. Ever since its development, the FIC has been increasingly applied in the realm of statistics, but this concept appears to be virtually unknown in the literature on energy and production economics. Using the classical example of the Translog cost function and production data for 35 U.S. industry sectors (1960–2005), this paper provides for an empirical illustration of the FIC and demonstrates its usefulness in selecting production models, thereby focusing on the ease of substitution between energy and capital versus energy and labor.

1. Introduction

There is an impressive bulk of empirical studies that aim at measuring the ease of substitution between production factors (for surveys, see e.g. Frondel and Schmidt, 2003; Kintis and Panas, 1989), with a growing emphasis on the substitution relationships of energy with respect to other inputs (see e.g. Apostolakis, 1990; Frondel and Schmidt, 2002, 2004). Common to the overwhelming majority of these studies is that the substitution parameters of interest are gleaned from a single 'best' production model whose selection may only be tenuously related to the purpose of inference.

There are numerous model selection methods, including the usage of information criteria, such as Akaike's (1974) AIC and Schwarz' (1978) SIC. Alternatively, Dette (1999), Dette et al. (2006), or Podolskij and Dette (2008) propose, among many others, goodness-of-fit tests. Typically, the selection of the most appropriate production model focuses on a few well-established functional forms, such as the Generalized Leontief and, most often, the Translog cost function. In seeking the right functional form, however, one might forget that any parametric model represents a highly stylized description of the real production process. As a consequence, none of these functional forms can claim to be the true model. Instead, depending on the facet of reality that is the focus of the analysis, divergent specifications might approximate different facets in an optimal way.

Recognizing this argument, Claeskens and Hjort (2003) deviated from the conventional avenue and conceived the Focused Information Criterion (FIC) to allow various models to be selected for different purposes. For instance, in the important case of the estimation of the degree of substitutability of energy and capital versus energy and labor, one kind of production model might be highly appropriate for inferences on, say, the cross-price elasticity of capital with respect to energy prices, whereas a different sort of model may be preferable for the estimation of another focus parameter, such as the cross-price elasticity of labor with respect to energy prices. In energy economics, such substitution elasticities are typical focus parameters that are important for evaluating a host of policy instruments, including, among other things, fuel and carbon taxes, as well as other environmental policy measures, such as energy efficiency standards.

Because of its usefulness in balancing modeling bias against estimation variability, the FIC has been increasingly applied in the realm of statistics (see e.g. Claeskens and Hjort, 2008; Claeskens et al., 2007; Hjort and Claeskens, 2006), but this concept appears to be virtually unknown in the economics literature, particularly in energy and production economics. The theoretical contribution of Behl et al. (2012) represents the sole exception for the literature on

* Corresponding author at: Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI), Hohenzollernstr. 1-3, D-45128 Essen, Germany.
E-mail address: frondel@rwi-essen.de (M. Frondel).

0140-9883/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.eneco.2013.04.013

economic modeling, while the analysis of Brownlees and Gallo (2008) is a rare example originating from financial economics.

Using the classical example of the Translog cost function and production data for 35 U.S. industry sectors in the time period spanning 1960 to 2005, this article provides for an empirical illustration of the usefulness of the FIC, demonstrating that the selection of a model type critically depends on the purpose of inference. It will become evident from our empirical example that model selection is highly dependent on the focus parameter μ, for instance whether the cross-price elasticity for either labor or capital demand with respect to energy prices is the primary aim of the analysis.

The general idea underlying the FIC, which ultimately results from estimating the mean squared error of the modeling bias (Claeskens and Hjort, 2003:902), is to study perturbations of a parametric model, with the known parameter vector $\gamma^0 := (\gamma_1^0,\dots,\gamma_q^0)^T$ as the point of departure. A variety of models may then be considered that depart from γ^0 in some or all of q directions: $\gamma \neq \gamma^0$. On the basis of parameter estimates of the altogether 2^q sub-models, that candidate model for which the FIC is minimal for a given focus parameter of choice $\mu = \mu(\gamma)$ will be selected.

By minimizing the FIC, one captures the trade-off between modeling bias, which, by definition, is zero for the most general model for which $\gamma_i \neq \gamma_i^0$ for $i = 1,\dots,q$, and relative estimation variability, which, by definition, is zero for the most restricted model for which $\gamma_i = \gamma_i^0$ for $i = 1,\dots,q$. In our empirical example, we deliberately confine ourselves to two polar model specifications, the Translog cost function as the most general specification and the Cobb–Douglas function as the most restricted of the 2^q model specifications, rather than estimating all of them. In fact, merely these polar specifications are of relevance in the economic literature, and, hence, bear special names, whereas the economic meaning of all other sub-models is minor.

The following Section 2 presents our example and derives the analytical expressions needed for the model selection among the Cobb–Douglas and Translog cost function on the basis of the FIC. Section 3 provides for a concise introduction into the concept of the FIC, followed by the presentation of the empirical example in Section 4. In Section 5, it will be explained that a comparison of the complete set of sub-models, rather than just the two polar model specifications, represents no principle difficulty. The last section summarizes and concludes.

2. Analytical example

To illustrate the concept of the FIC on the basis of a straightforward example that is – for the sake of simplicity – restricted to the case of three production factors, we employ the dual approach (Berndt, 1996), in which a system of cost share equations is derived from the underlying cost function via Shepard's lemma. For a Cobb–Douglas cost function, cost shares are well-known to be independent from factor prices:

$$s_K = \beta_K + \nu_K, \quad s_L = \beta_L + \nu_L, \tag{1}$$

where s_K and s_L denote the cost shares of capital K and labor L, respectively, and β_L and β_K are parameters to be estimated. As Maximum Likelihood (ML) is Claeskens' and Hjort's (2003) estimation method of choice, we assume joint normality for the random errors v_L and v_K: $(\nu_K,\nu_L) \sim N((0,0),(\sigma_K^2,\sigma_L^2),\rho_{KL})$, where σ_K^2 and σ_L^2 designate the variances of v_L and v_K, respectively, and ρ_{KL} stands for the correlation of the error terms: $\rho_{KL} := Corr(\nu_K,\nu_L)$. Adding a third equation for the cost share of energy E to system (1) would be superfluous, as the cost shares sum up to unity: $s_K + s_L + s_E = 1$. Implicitly, this property yields the restrictions $\beta_K + \beta_L + \beta_E = 1$ and $\nu_K + \nu_L + \nu_E = 0$, so that an estimate of $\beta_E = 1 - \beta_L - \beta_K$ can be obtained from the estimates of β_L and β_K.

For the same reason, it suffices to estimate the following two-equation system for the Translog cost function[1]:

$$s_K = \beta_K + \beta_{KK} p_K + \beta_{KE} p_E + \nu_K, \quad s_L = \beta_L + \beta_{LK} p_K + \beta_{LE} p_E + \nu_L, \tag{2}$$

where p_K and p_E denote the logged relative factor prices $log(\tilde{p}_K/\tilde{p}_L)$ and $log(\tilde{p}_E/\tilde{p}_L)$, respectively. If $\beta_{KK} = \beta_{KE} = \beta_{LK} = \beta_{LE} = 0$, the Translog specification degenerates to the Cobb–Douglas case.

Adopting the terminology of Claeskens and Hjort (2003), the Cobb–Douglas specification (1) is called the null model. For this specification, also referred to as the narrow model, the vector ξ of parameters that are subject to estimation comprises four elements:

$$\xi := (\beta_K, \beta_L, \sigma_K, \sigma_L)^T, \tag{3}$$

where T indicates the transposition of a vector and σ_K and σ_L designate the standard deviations of v_L and v_K, respectively. The vector of parameters that are additionally included in the Translog model, which is called the full model, reads:

$$\gamma := (\beta_{KK}, \beta_{KE}, \beta_{LK}, \beta_{LE}, \rho_{KL})^T. \tag{4}$$

For clarity, the parameters of the null model are denoted by $\theta^0 := (\xi^0,\gamma^0)^T$, with $\xi^0 := (\beta_K^0,\beta_L^0,\sigma_K^0,\sigma_L^0)^T$ and $\gamma^0 = 0$ if we additionally assume $\rho_{KL} = 0$ for the Cobb–Douglas case.

In contrast to conventional selection criteria, using the FIC for model selection orients towards one or more measures of interest, called here focus parameters and designated by μ, which are typically a function of the model coefficients: $\mu = \mu(\xi,\gamma)$. As our focus is on the substitutability of energy by both labor and capital, we choose the cross-price elasticities of capital and labor demand, both with respect to the price of energy, as focus parameters. For the Translog model (2), these substitution elasticities are given by (see e.g. Frondel and Schmidt, 2006)[2]:

$$\eta_{K_{p_E}} = \frac{\beta_{KE}}{s_K} + s_E = \frac{\beta_{KE}}{s_K} + 1 - s_K - s_L, \tag{5}$$

$$\eta_{L_{p_E}} = \frac{\beta_{LE}}{s_L} + s_E = \frac{\beta_{LE}}{s_L} + 1 - s_K - s_L, \tag{6}$$

where according to system (2) the cost shares of capital and labor itself depend on coefficients such as β_K, β_{KK}, etc. and are stochastic, inheriting this property to the cross-price elasticities. Yet, in what follows, to keep the exposition simple, we ignore the stochastic nature of the cross-price elasticities, treating them as deterministic. This approach is in line with the empirical estimation of these elasticities in Section 4, for which we use the estimated cost shares, rather than the observed cost shares.

As we will see in the subsequent section, the dependence of the FIC on a focus measure μ – here the elasticities $\eta_{L_{p_E}}$ and $\eta_{K_{p_E}}$ – is given by the vectors of partial derivatives of such measures with

[1] The Translog cost function underlying cost share system (2) reads as follows:

$$\begin{aligned} logC = \beta_0 &+ \beta_K log\,\tilde{p}_K + \beta_L log\,\tilde{p}_L + \beta_E log\,\tilde{p}_E \\ &+ \frac{1}{2}\beta_{KK}(log\,\tilde{p}_K)^2 + \beta_{LK} log\,\tilde{p}_K log\,\tilde{p}_L + \beta_{KE} log\,\tilde{p}_K log\,\tilde{p}_E \\ &+ \frac{1}{2}\beta_{LL}(log\,\tilde{p}_L)^2 + \beta_{LE} log\,\tilde{p}_L log\,\tilde{p}_E + \frac{1}{2}\beta_{EE}(log\,\tilde{p}_E)^2, \end{aligned}$$

where log denotes the natural logarithm. For $\beta_{KK} = \beta_{LK} = \beta_{KE} = \beta_{LL} = \beta_{LE} = \beta_{EE} = 0$, this Translog cost function degenerates to the COBB–DOUGLAS cost function and equation system (2) reduces to system (1).

[2] In empirical contexts, the cross-price elasticities are well-defined, because in practice the cost shares of production factors are always larger than zero, meaning that these factors are economically relevant. Cost shares may be quite close to zero, though, specifically those of the production factor energy.