Helsinki
Center
of
Economic
Research

Discussion Papers

A Critique of the System Estimation Approach of Normalized CES Production Functions

Arto Luoma
University of Tampere

and

Jani Luoto
University of Helsinki and HECER

Discussion Paper No. 336
September 2011

ISSN 1795-0562

HECER – Helsinki Center of Economic Research, P.O. Box 17 (Arkadiankatu 7), FI-00014 University of Helsinki, FINLAND, Tel +358-9-191-28780, Fax +358-9-191-28781,
E-mail info-hecer@helsinki.fi, Internet www.hecer.fi

HECER
Discussion Paper No. 336

A Critique of the System Estimation Approach of Normalized CES Production Functions*

Abstract

Recent evidence indicates that the supply side system approach which models the CES production function and profit maximizing conditions jointly is superior to other empirical approaches. This evidence is mainly based on a nonlinear feasible generalized least squares (FGLS) estimation method. We show that in order to employ the nonlinear FGLS method, contradictory assumptions about the errors of the supply side equations are required. We propose a Bayesian full information method which, by construction, avoids such inconsistency. A local identification problem appearing in the analysis of the CES production function under nonneutral technical change is further pointed out.

JEL Classification: E22, O33, O41, C11, C30

Keywords: constant elasticity of substitution, factor-augmenting technical change, simultaneous equations, Bayesian analysis.

Arto Luoma

Department of Mathematics and Statistics
University of Tampere
FI-33014 University of Tampere
FINLAND

e-mail: arto.luoma@uta.fi

Jani Luoto

Department of Political and Economic Studies
University of Helsinki
P.O. Box 17 (Arkadiankatu 7),
FI-00014 University of Helsinki
FINLAND

e-mail: jani.luoto@helsinki.fi

* We would like to thank Bob Chirinko and Markku Lanne for useful comments. The usual disclaimer applies. Financial support from the Academy of Finland, the OP-Pohjola Group Research Foundation and the Yrjö Jahnsson Foundation is gratefully acknowledged. This paper was written while Luoto was visiting the Faculty of Economics of the University of Cambridge.

1. Introduction

In a series of articles Klump et al. (2007a, 2007b, 2008) and León-Ledesma et al. (2010a) asserted that the supply side system approach which models the normalized[1] CES production function and profit maximizing conditions jointly is superior to other empirical approaches (see Chirinko, 2008, for survey). Their evidence indicates that the success of the system approach lies in its ability to jointly identify the substitution elasticity, σ, and the nonneutral technical change parameters. We argue that despite this advantage their estimation approach based on a nonlinear feasible generalized least squares (FGLS) method is inappropriate because it is internally inconsistent. Specifically, we show that if the errors of the supply side equations are correlated, then the system incorporates instantaneous feedback effects into these equations.[2] The problem is that the FGLS estimator is particularly designed for accounting for possible correlations of these errors, but it is valid only when the estimated equations are exogenous. Thus we see that the estimation approach of Klump et al. (2007a, 2007b, 2008) and León-Ledesma et al. (2010a) conflicts with itself. We instead propose a Bayesian full information method for the system estimation, which, by construction, avoids such inconsistency.[3]

Our results, based on a Monte Carlo (MC) simulation study borrowed from León-Ledesma et al. (2010a), show that the CES production function and the first-order conditions of profit maximization indeed form a (nonlinear) system of simultaneous equations. For plausible substitution elasticity values, our method corrects implied simultaneity bias and accurately identifies the substitution elasticity parameter and technical progress parameters. However, if the simultaneous equations problem is not solved, the estimates of the substitution elasticity parameter are systematically biased towards unity.

[1] The necessity of normalizing the CES production function to obtain clear and unambiguous interpretations for its parameters was first pointed out by de La Grandville (1989) and further studied by Klump and de La Grandville (2000) and Klump and Preissler (2000) (see Klump et al., 2011; and references therein). Its usefulness in empirical analysis, especially in the system case, is shown by Klump et al. (2007a, 2007b, 2008) and León-Ledesma et al. (2010a, b).

[2] The authors also checked the performance of the system method by using a 3SLS estimator. The obvious problem with instrumental variables (IV) estimators in the case of CES production functions is the lack of good instruments (e.g., Chirinko, 2008, and León-Ledesma et al., 2010a), resulting in weak identification and strong dependence of results on the choice of instruments.

[3] Our proposal to estimate the CES production function from the Bayesian viewpoint has its roots in the literature of direct estimation of nonlinear CES production functions. The idea of estimating the parameters of a nonlinear CES function directly was introduced by Chetty and Sankar (1969). Since then, it has become a standard text book example in Bayesian econometrics (see, e.g., Zellner, 1971, Koop, 2003, and Lancaster, 2004).

In addition to the simultaneous equations problem, our simulation study reveals another empirically important issue hampering the estimation of the CES production function with nonneutral technical change. Namely, technical change parameters are not identifiable under Cobb-Douglas production. This local identification issue arises from the well-known fact that when σ = 1, the technical parameters enter neither the CES production function nor the first-order conditions of profit maximization. As a result, the supply side system is non-informative in the dimensions of the technical parameters when $\sigma \approx 1$, which tends to hamper the system estimation of the substitution elasticity and technical change parameters.

The rest of the paper is structured as follows. In Section 2, we describe the normalized supply side system approach and discuss related issues. Section 3 provides a method for Bayesian estimation. Section 4 presents Monte Carlo results. Finally, Section 5 concludes.

2. Normalized System Approach

In this section, we first describe the normalized supply side system approach proposed by Klump et al. (2007a, 2007b, 2008) and León-Ledesma et al. (2010a), and then show that it is internally inconsistent when based on the nonlinear FGLS method. We further point out the local identification issue already discussed in Introduction.

In recent years, there has been increasing interest toward the joint estimation of the elasticity of substitution between capital and labor and the direction of technical change parameters. In this context, the research often draws on the so called normalized CES production function in order to obtain a direct economic interpretation for its parameters. The form of the normalized CES production function which allows for biased technical change is

$$\frac{Y_t}{Y_0} = \left[\alpha_0 \left(\frac{\Gamma_t^K K_t}{\Gamma_0^K K_0} \right)^{\frac{\sigma-1}{\sigma}} + (1-\alpha_0) \left(\frac{\Gamma_t^L L_t}{\Gamma_0^L L_0} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \tag{1}$$

where Y_t is the real output, K_t is the real capital stock, L_t is the labour input, and α_0 is the capital income share at the point of normalization t = 0. The elasticity of substitution between capital and labour inputs can be expressed as

$$\sigma = \frac{d\log(K/L)}{d\log(F_L/F_K)}, \tag{2}$$

where F_i is the marginal product of factor i. The capital- and labour-augmenting technical progress is captured by Γ_t^K and Γ_t^L. We follow the standard practice in assuming that technical progress is linear. In particular, its exact form is provided by $\Gamma_t^K = e^{\gamma_K (t-t_0)}$ and $\Gamma_t^L = e^{\gamma_L (t-t_0)}$, where γ_i is the growth rate of the technical change associated to factor i.[4]

If the factors are paid according to their marginal products, then from (1) we obtain the standard first order conditions (FOC) of profit maximization:

$$F_K = \frac{\partial Y_t}{\partial K_t} = \alpha_0 \left(\frac{\Gamma_t^K Y_0}{\Gamma_0^K K_0} \right)^{\frac{\sigma-1}{\sigma}} \left(\frac{Y_t}{K_t} \right)^{\frac{1}{\sigma}} = r_t, \tag{3}$$

$$F_L = \frac{\partial Y_t}{\partial L_t} = (1-\alpha_0) \left(\frac{\Gamma_t^L Y_0}{\Gamma_0^L L_0} \right)^{\frac{\sigma-1}{\sigma}} \left(\frac{Y_t}{L_t} \right)^{\frac{1}{\sigma}} = w_t, \tag{4}$$

where r_t and w_t are the real interest rate and real wage, respectively. The normalized supply side system proposed by Klump et al. (2007a, 2007b, 2008) and León-Ledesma et al. (2010a) can be obtained by taking logs from (3), (4) and (1), and using the sample averages (geometric for Y_0, K_0 and L_0, and arithmetic for $\alpha_0 = r_0 K_0 / Y_0$ and t_0) as the point of normalization. The system is given by

$$\ln(r_t) = \ln\left(\frac{\overline{\alpha}\overline{Y}}{\overline{K}} \right) + \frac{\sigma-1}{\sigma}\ln(\vartheta) + \frac{1}{\sigma}\ln\left(\frac{Y_t/\overline{Y}}{K_t/\overline{K}} \right) + \frac{\sigma-1}{\sigma}\gamma_K \tilde{t} + \varepsilon_{1t}, \tag{5}$$

$$\ln(w_t) = \ln\left(\frac{(1-\overline{\alpha})\overline{Y}}{\overline{L}} \right) + \frac{\sigma-1}{\sigma}\ln(\vartheta) + \frac{1}{\sigma}\ln\left(\frac{Y_t/\overline{Y}}{L_t/\overline{L}} \right) + \frac{\sigma-1}{\sigma}\gamma_L \tilde{t} + \varepsilon_{2t}, \tag{6}$$

[4] Alternative forms of technical progress may be considered. For example, Klump et al. (2007a,b) and León-Ledesma et al. (2010a), propose a more flexible functional form for Γ_t^i based on the Box-Cox transformation, in order to capture the potentially nonconstant rates of technical progress. We do not consider issues raised by induced innovations, and, thus, linear technical progress works well for our purpose.

$$\ln\left(\frac{Y_t}{\overline{Y}}\right)=\ln(\vartheta)+\ln\left[\overline{\alpha}\left(\frac{K_t}{\overline{K}}e^{\gamma_K\tilde{t}}\right)^{\frac{\sigma-1}{\sigma}}+\left(1-\overline{\alpha}\right)\left(\frac{L_t}{\overline{L}}e^{\gamma_L\tilde{t}}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}+\varepsilon_{3t}\,, \tag{7}$$

where a bar refers to the sample average of the variable in question, $\tilde{t}=t-\bar{t}$, ϑ is an extra term capturing the deviation of $\overline{Y}$ from the level of production at the sample averages of the right-hand side variables of (1), and ε_{1t}, ε_{2t}, and ε_{3t} are stochastic shocks temporarily deviating r_t, w_t, and Y_t from their equilibrium values. As already discussed, the nonlinear FGLS method, which assumes exogenous right hand side variables and allows for possible correlations between the shocks ε_{1t}, ε_{2t}, and ε_{3t}, has been the main estimation method applied for the system (5)-(7). In the following discussion, we will show that these two assumptions contradict, implying that the estimation approach of Klump et al. (2007a, 2007b, 2008) and León-Ledesma et al. (2010a) is inappropriate.

We assume that the vectors $\boldsymbol{\varepsilon}_t=(\varepsilon_{1t},\varepsilon_{2t},\varepsilon_{3t})'$ are independent and identically distributed (i.i.d.) multivariate normal:

$$\boldsymbol{\varepsilon}_t\sim N(\mathbf{0},\Sigma),\ \text{where}\ \Sigma=\begin{pmatrix}\sigma_{11}&\sigma_{12}&\sigma_{13}\\\sigma_{21}&\sigma_{22}&\sigma_{23}\\\sigma_{31}&\sigma_{32}&\sigma_{33}\end{pmatrix}. \tag{8}$$

A few comments about this assumption are in order. The normal distribution is quite reasonable for ε_{1t} and ε_{2t} due to the log-linear form of (3) and (4). Of course, the i.i.d. assumption is based on fully flexible prices (that is, economic cycles are abstracted), whereas the real-world real interest rate and real wage series consist of strong permanent components. However, our results are not sensitive to the autocorrelation of the errors. As far as ε_{3t} is concerned, its distribution is not so clear-cut. To see this, notice that under flexible prices ε_{3t} should be interpreted as a measurement error arising from the (log) difference between the observed and the potential (equilibrium) output (given in the right hand side of (1)). In absence of the pure profit component the observed output is obtained from the accounting identity $Y_t\equiv r_tK_t+w_tL_t$. It can be shown, using this identity and (5) and (6), that the exponential of ε_{3t} is a weighted average of two log-normally distributed random variables, namely, the exponentials of ε_{1t} and ε_{2t} (see León-Ledesma et al., 2010a). Thus, if r_t and w_t are as assumed in (5) and (6), then ε_{3t} is not Gaussian, but could be adequately approximated by a Gaussian distribution. This shock structure also implies that the same shocks which drive r_t and w_t also drive

Y_t, which appears on the right hand side of the system. This, in turn, implies that (5)-(7) form a system of simultaneous equations. Therefore, from now on, it is referred to as the structural form system.

The reduced form supply side equations can be obtained from (5)-(8) in a standard fashion. They can be expressed as

$$\ln(r_t) = \ln\left(\frac{\overline{\alpha}\overline{Y}\vartheta}{\overline{K}}\right) + \frac{1}{\sigma}\ln\left(\frac{Y_t^*(\sigma, \gamma_K, \gamma_L)}{K_t/\overline{K}}\right) + \frac{\sigma - 1}{\sigma}\gamma_K \tilde{t} + e_{1t}\,, \tag{9}$$

$$\ln(w_t) = \ln\left(\frac{(1-\overline{\alpha})\overline{Y}\vartheta}{\overline{L}}\right) + \frac{1}{\sigma}\ln\left(\frac{Y_t^*(\sigma, \gamma_K, \gamma_L)}{L_t/\overline{L}}\right) + \frac{\sigma - 1}{\sigma}\gamma_L \tilde{t} + e_{2t}\,, \tag{10}$$

$$\ln\left(\frac{Y_t}{\overline{Y}}\right) = \ln(\vartheta) + \ln\left(Y_t^*(\sigma, \gamma_K, \gamma_L)\right) + e_{3t}\,, \tag{11}$$

where $Y_t^*(\sigma, \gamma_K, \gamma_L)$ is such that (7) and (11) are identical, and the new error vector $\mathbf{e}_t = (e_{1t}, e_{2t}, e_{3t})'$ is given by

$$\mathbf{e}_t = \Gamma^{-1}\boldsymbol{\varepsilon}_t \sim N(\mathbf{0},\Psi),\ \Psi = \Gamma^{-1}\Sigma(\Gamma^{-1})',\ \text{where}\ \Gamma^{-1} = \begin{pmatrix} 1 & 0 & 1/\sigma \\ 0 & 1 & 1/\sigma \\ 0 & 0 & 1 \end{pmatrix}. \tag{12}$$

Notice that the system of equations (9)-(12) determines the distribution of the data vector $\mathbf{x}_t$ = $(\ln(r_t), \ln(w_t), \ln(Y_t/\overline{Y}))'$ for $t = 1,\ldots,T$, giving rise to the likelihood function

$$L(X; \theta, K, L) \propto |\Psi|^{-T/2} \exp\left\{-\frac{1}{2}tr\left(E'E\,\Psi^{-1}\right)\right\}, \tag{13}$$

where tr denotes the trace of a matrix, $\theta = (\sigma, □, \gamma_K, \gamma_L, \psi_{11}, \psi_{12}, \psi_{13}, \psi_{22}, \psi_{23}, \psi_{33})'$ is a vector containing all the parameters of the model, ψ_{ij} are the elements of Ψ, $K = (K_1,\ldots, K_T)'$, $L = (L_1,\ldots,$

$L_T)'$, and the matrices X and E are obtained by stacking the row vectors $\mathbf{x}_t'$ and $\mathbf{e}_t'$, respectively, for $t = 1,\ldots T$.

Let us next consider the conditional distribution of $\mathbf{x}_{1t} = (\ln(r_t), \ln(w_t))'$ given $\mathbf{x}_{2t} = \ln\left(Y_t/\overline{Y}\right)$. From the system of equations (9)-(12), by using the properties of the multivariate normal distribution and the fact that $\Psi_{12} = (\sigma_{13}+\sigma_{33}/\sigma, \sigma_{23}+\sigma_{33}/\sigma)'$, where Ψ is partitioned as

$$\Psi = \begin{pmatrix} \Psi_{11} & \Psi_{12} \\ \Psi'_{12} & \psi_{33} \end{pmatrix}, \ \psi_{33} = \sigma_{33},$$

we obtain

$$\ln(r_t) = \mu_1 + \frac{1}{\sigma}\ln\left(\frac{Y_t/\overline{Y}}{K_t/\overline{K}}\right) + \frac{\sigma-1}{\sigma}\gamma_K\tilde{t} + \frac{\sigma_{13}}{\sigma_{33}}\ln\left(\frac{Y_t/\overline{Y}}{Y_t^*(\sigma,\gamma_K,\gamma_L)}\right) + \eta_{1t}, \tag{14}$$

$$\ln(w_t) = \mu_2 + \frac{1}{\sigma}\ln\left(\frac{Y_t/\overline{Y}}{L_t/\overline{L}}\right) + \frac{\sigma-1}{\sigma}\gamma_L\tilde{t} + \frac{\sigma_{23}}{\sigma_{33}}\ln\left(\frac{Y_t/\overline{Y}}{Y_t^*(\sigma,\gamma_K,\gamma_L)}\right) + \eta_{2t}, \tag{15}$$

where

$$\mu_1 = \ln\left(\overline{\alpha}\,\overline{Y}/\overline{K}\right) + [(\sigma-1)/\sigma - (\sigma_{13}/\sigma_{33})]\ln(\square),$$

$$\mu_2 = \ln\left((1-\overline{\alpha})\overline{Y}/\overline{L}\right) + [(\sigma-1)/\sigma - (\sigma_{23}/\sigma_{33})]\ln(\square),$$

$$\boldsymbol{\eta}_t = \begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix} \sim N(\mathbf{0},\Omega),$$

$$\Omega = \Psi_{11} - \Psi_{12}\Psi'_{12}/\psi_{33}.$$

Equations (14)-(15) together with (7) (or with (11)) provide an equivalent representation for the reduced form system in (9)-(11). Therefore, by working out with the basic rules of conditional probability, the likelihood function in (13) can be equivalently expressed as

$$L(X;\theta,K,L) = L_1(X_1;\theta,X_2,K,L)L_2(X_2;\theta,K,L), \tag{16}$$

where

$$L_1(X_1;\theta,X_2,K,L) \propto |\Omega|^{-T/2} \exp\left\{-\frac{1}{2} tr \mathrm{H}'\Omega^{-1}\mathrm{H}\right\}, \tag{17}$$

$$L_2(X_2;\theta,K,L) \propto \sigma_{33}^{-T/2} \exp\left\{-\frac{1}{2}\sum_{t=1}^{T}(\varepsilon_{3t})^2 \Big/ \sigma_{33}\right\}, \tag{18}$$

$X_2 = (\mathbf{x}_{21},\ldots, \mathbf{x}_{2T})'$, and the *t*ht rows of the matrices X_1 and H are given by $\mathbf{x}_{1t}'$ and $\boldsymbol{\eta}_t'$, respectively. Thus, by comparing (5)-(7) to (14)-(16) it can be seen that the structural and reduced form systems coincide under the assumption $\sigma_{12} = \sigma_{13} = \sigma_{23} = 0$. However, if the underlying economic system is consistent with national-accounting practice, then the shock ε_{3t} is driven by ε_{1t} and ε_{2t} (that is, $\sigma_{13} \neq 0$ and $\sigma_{23} \neq 0$) and, as a result, the right hand side variable Y_t in (5) and (6) is correlated with ε_{1t} and ε_{2t}. In Equation (16) these contemporaneous relationships are captured by the second last terms on the right hand sides of (14) and (15). If we were to estimate the system of equations (5)-(7) using, for example, the nonlinear least squares method, we would neglect the information provided by these terms (i.e., we would impose the restriction $\sigma_{13} = \sigma_{23} = 0$ into the system). It is well-known from linear simultaneous equations models that this would lead to biased parameter estimates.

As already discussed, Klump et al. (2007a, 2007b, 2008) and León-Ledesma et al. (2010a) estimate the structural form system (5)-(7) using the nonlinear FGLS method, which takes into account the potential correlations of shocks by estimating Σ but neglects the information provided by the second last terms of (14) and (15). The problem is that they estimate Σ freely, but at the same time implicitly set the elements σ_{13} and σ_{23} at zero (cf., (14) and (15)). This internal inconsistency casts serious doubts on the reliability of their estimation approach. From an econometric standpoint their method fails to distinguish between the distributions of the shocks and the data. From an economic point of view their system does not take into account the information provided by the accounting identity $Y_t \equiv r_tK_t + w_tL_t$, which links the factor prices and output as a (nonlinear) system of simultaneous equations.

In order to avoid the inconsistency discussed above we propose estimating the system using a Bayesian method, based on the likelihood in (13). As can be seen from (9)-(11), the method treats the potential output $Y_t^*(\sigma,\gamma_K,\gamma_L)$ as an instrument for the observed output. Thus, it has the advantage that no instrumental variables are needed. The obvious problem with instrumental variables (IV) estimators in our case is the lack of good instruments (see, e.g., Chirinko, 2008, and León-Ledesma et al., 2010a), resulting in weak identification and strong dependence of results on the choice of instruments.

An important issue related to the empirical analysis of (13) is that of local non-identification. In particular, the technical change parameters γ_K and γ_L are not identified at $\sigma = 1$ (i.e., in the Cobb-Douglas case). The reason is that when $\sigma = 1$, the parameters γ_K and γ_L do not enter the likelihood function, as can be seen from Equations (5)-(7) (cf., the fact that $(\sigma–1)/\sigma = 0$ at $\sigma = 1$). Thus, the likelihood function is flat (and the system non-informative) in the dimensions of γ_K and γ_L when $\sigma \approx 1$.[5]

3. Alternative Bayesian Approach

We shall now consider the full conditional distributions of Ψ, $\xi = \ln(\square)$, γ_K, γ_L, and σ, which can be used to estimate these parameters. Furthermore, we describe the adopted joint prior distribution.

For the Bayesian analysis we need to specify the prior distribution of the parameters, in addition to the likelihood (13). We shall assume that

$$\begin{aligned} p\left(\sigma,\gamma_K,\gamma_L,\xi,\Psi^{-1}\right) &= p(\sigma)p(\gamma_K)p(\gamma_L)p(\xi)p\left(\Psi^{-1}\right) \\ &\propto N\left(\gamma_K\middle|\tilde{\gamma}_K,\kappa_K^2\right)N\left(\gamma_L\middle|\tilde{\gamma}_L,\kappa_L^2\right)N\left(\xi\middle|\tilde{\xi},\tau^2\right) \\ &\quad\cdot W\left(\Psi^{-1}\middle|\nu,\Psi_0\right)I_{(0,\infty)}(\sigma), \end{aligned} \tag{19}$$

where $I_{(0,\infty)}(\sigma)$ is an indicator function obtaining value one for positive σ and zero otherwise, $W(\nu,\Psi_0)$ refers to a Wishart distribution with inverse scale matrix Ψ_0 and degrees of freedom parameter ν, and $\tilde{\xi}$, $\tilde{\gamma}_L$, $\tilde{\gamma}_K$, τ, κ_K and κ_L are the remaining prior hyper parameters. Under this joint prior, the full conditional distributions of Ψ^{-1}, $\xi = \ln(\square)$, γ_K, γ_L, and σ are given by

[5] Notice that in the Bayesian approach we need to combine the likelihood function with a prior distribution of parameters in order to make inference. If the conditional prior distribution of γ_K and γ_L (at $\sigma = 1$) is improper (i.e., does not integrate to one), then, as a result of the flat likelihood, the posterior distribution is also improper. As a potential solution for the same type of problem, existing in the Bayesian analysis of unrestricted simultaneous equations models, Kleibergen and van Dijk (1998) propose using informative priors. This solution is also adopted here, and we shall return to this issue in Section 3. The reader may also be interested to know that the same type of local non-identification issue also exists in Bayesian cointegration analysis (see Koop et al., 2005, for survey).

$$\Psi^{-1}\big|\,\xi,\sigma,\gamma_K,\gamma_L \sim W\left(T+v,\left(\Psi_0^{-1}+S\right)^{-1}\right), \tag{20}$$

$$p\left(\xi\middle|\Psi,\sigma,\gamma_K,\gamma_L\right) \propto \exp\left\{-\frac{1}{2}tr\left(S\,\Psi^{-1}\right)\right\}\exp\left\{-\frac{1}{2\tau^2}\left(\xi-\tilde{\xi}\right)^2\right\}, \tag{21}$$

$$p\left(\gamma_K\middle|\Psi,\xi,\sigma,\gamma_L\right) \propto \exp\left\{-\frac{1}{2}tr\left(S\,\Psi^{-1}\right)\right\}\exp\left\{-\frac{1}{2\,\kappa_K^2}\left(\gamma_K-\tilde{\gamma}_K\right)^2\right\}, \tag{22}$$

$$p\left(\gamma_L\middle|\Psi,\xi,\sigma,\gamma_K\right) \propto \exp\left\{-\frac{1}{2}tr\left(S\,\Psi^{-1}\right)\right\}\exp\left\{-\frac{1}{2\kappa_L^2}\left(\gamma_L-\tilde{\gamma}_L\right)^2\right\}, \tag{23}$$

$$p\left(\sigma\middle|\Psi,\xi,\gamma_K,\gamma_L\right) \propto \exp\left\{-\frac{1}{2}tr\left(S\,\Psi^{-1}\right)\right\}I_{(0,\infty)}(\sigma), \tag{24}$$

where $S = E'E$, E being the same as in (13), and Σ can be calculated by mapping from Ψ to Σ, given Γ.[6] Equation (20) can be derived by combining the natural conjugate Wishart prior $W(v,\Psi_0)$ with the likelihood in (13) and using standard calculations (e.g., Koop, 2003), while equations (21)-(24) follow directly from the definition of the full conditional distribution.

Our sampler involves sequential drawings from the full conditional distributions (20)-(24).[7] Our experience is that convergence occurs rapidly when σ is not too close to unity, without further tuning of the sampler. The sampler may be inefficient if σ is close to unity, producing highly correlated chains (because of the local non-identification problem discussed previously). This also indicates that the model is misspecified, since under the Cobb-Douglas case technological progress degenerates to the Hicks-neutral case ($\gamma_K = \gamma_L > 0$). In practice, in near unitary substitution the

[6] It can be seen from (12) that the one-to-one mapping between Σ and Ψ is unique, given σ. Because σ can be obtained independently of Ψ (e.g., by analytically integrating Ψ out from (13) and then using the resulting marginal distribution to obtain the rest of the parameters), this "conditional uniqueness" of the one-to-one mapping ensures that the model is globally identified.

[7] The full conditional posterior distribution given in (20) is standard and can be readily used to simulate random numbers. For the rest of the parameters ln(□), σ, γ_K, and γ_L, however, the situation is more intricate. Fortunately, suitable candidate-generating densities, which provide high acceptance rates for candidate draws, are available for these parameters. In particular, a univariate normal distribution, with mean at the mode of one of the conditional posteriors (21)-(24) and precision equal to the negative of the second derivative of the log posterior, evaluated at the mode, can successfully be used as a candidate distribution for each ln(□), σ, γ_K, and γ_L, and the acceptance-reception probability can be calculated in the standard way.

supply side model is over-parameterized and unable to capture the trend in Y_t (cf., (7)). Thus, we recommend not using the system if the posterior distribution of σ lies near unity.

A natural default choice for the prior hyper parameters Ψ_0 and v would be such that the prior for Ψ is as noninformative as possible. Noninformativeness is achieved by setting $\Psi_0^{-1} = \mathbf{0}$ and $v = 0$. On the other hand, if $\sigma_{ij} = 0$ for $i \neq j$, the model is over-parameterized and an informative prior might be a good choice. As a compromise, the degrees of freedom parameter v and the diagonal elements of the scale matrix Ψ_0^{-1} are set to be small, namely 5 and 0.01, respectively. In addition, the non-diagonal elements of Ψ_0^{-1} are set at zero. As discussed by León-Ledesma et al. (2010a), □ should be close to one, suggesting that the natural default choice for $\tilde{\xi}$ is zero. We render the prior of $\xi = \ln(□)$ noninfluential by setting $\tau = 100$.

Because of the local identification problem the prior variances of γ_K and γ_L cannot be too large. The basic idea is to choose κ_K and κ_L to be small enough in order to ensure that the technology parameters γ_K and γ_L cannot deviate too much from their true values under near unitary substitution. In practice, the choice of κ_K and κ_L should be such that it corresponds to a reasonable growth rate of technical progress. In this paper, the results are reported for $\kappa_K = \kappa_L = \{1/20, 1/10\}$. The value 1/10 yields rather noninformative priors for the technical (growth rate) parameters, while the choice 1/20 gives considerably more weight to the zero means ($\tilde{\gamma}_L = \tilde{\gamma}_K = 0$).[8]

4. Monte Carlo Analysis

We use simulated numerical examples based on a variety of parameter values to demonstrate the endogeneity and local identification issues discussed previously. For each example, we simulate $M = 1000$ samples of size $T = 50$[9] for the capital K_t, labour L_t, and technology functions Γ_t^K and Γ_t^L, and use these to calculate the potential and observed outputs and the real factor payments. For each sample, we then estimate the supply side system by simulating a Markov chain of $N = 3000$ cycles (abstracting 250 burn-in cycles) from the full conditional distributions explained in Section 3.[10] The median values of the Markov chains are used as point estimates. Two types of Bayesian system

[8] We also considered flatter priors for these parameters in our simulation experiments (when setting the true value of σ to be far from unity) but this had virtually no effect on the results.
[9] Using $T = 100$ leads to very similar results.
[10] The results based on $N = 10000$ cycles (abstracting 2000 burn-in cycles) seem to be practically the same, which indicates rapid convergence of the Markov chains (this was checked with a range of parameter values).

approaches are compared: (*i*) the unrestricted reduced form system approach introduced in Section 2; (*ii*) the structural form system approach which is obtained from the former by imposing the restriction $\sigma_{12} = \sigma_{13} = \sigma_{32} = 0$.[11] For both estimation methods the averages of the point estimates across the 1000 runs and the associated standard errors are reported. The distributions of these point estimates are further plotted, as they turned out to be quite informative.

4.1. Experiment

The simulation experiment is taken from León-Ledesma et al. (2010a). It includes draws from the stochastic labor $L_t = L_{t-1}\exp\left(l + \varepsilon_t^L\right)$, capital $K_t = K_{t-1}\exp\left(k + \varepsilon_t^K\right)$, and technical progress functions given by

$$\frac{\Gamma_t^L}{\Gamma_0^L} = \exp\left(\gamma_L\left(t - t_0\right) + \varepsilon_t^{\Gamma^L}\right),\ \frac{\Gamma_t^K}{\Gamma_0^K} = \exp\left(\gamma_K\left(t - t_0\right) + \varepsilon_t^{\Gamma^K}\right), \text{ for } t = 1,\ldots,T, \tag{25}$$

where l and k refer to the mean growth rates of labour supply and capital accumulation, respectively. Thus, it is assumed that the log production factors are random walks with drift, while capital- and labour-augmenting technical progress is driven by a deterministic trend plus stochastic technology shocks. Shocks to L_t, K_t, Γ_t^L, and Γ_t^K are assumed to be Gaussian, $\varepsilon_t^i \sim N\left(0,\sigma_i^2\right)$ and $\varepsilon_t^{\Gamma^i} \sim N\left(0,\sigma_{\Gamma^i}^2\right)$, for $i = L, K$, and the initial values are set at $L_0 = K_0 = \Gamma_0^L = \Gamma_0^K = 1$ for simplicity.

Given the simulated paths of L_t, K_t, Γ_t^L, and Γ_t^K, the next step is to calculate the output from the CES production function

$$Y_t^* = Y_0^*\left[\alpha_0\left(\frac{\Gamma_t^K K_t}{\Gamma_0^K K_0}\right)^{\frac{\sigma-1}{\sigma}} + \left(1-\alpha_0\right)\left(\frac{\Gamma_t^L L_t}{\Gamma_0^L L_0}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}, \tag{26}$$

[11] In this case case (20) reduces to three independent gamma distributions and the likelihood function in (16) (with the restriction $\sigma_{13} = \sigma_{32} = \sigma_{12} = 0$) can be employed instead of (13). We use a standard gamma prior for σ_{ii} with the shape and rate parameters set at 0.01 (thus, the priors are virtually noninfluential). For the rest of the parameters the priors explained in Section 3 are adopted.

where the star refers to the potential (equilibrium) output (cf., the reduced form system (9)-(11) in Section 2).

The formulas for the real factor payments r_t and w_t can be obtained from (26) using the standard FOCs:

$$r_t = \alpha_0 \left(\frac{\Gamma_t^K Y_0^*}{\Gamma_0^K K_0} \right)^{\frac{\sigma-1}{\sigma}} \left(\frac{Y_t^*}{K_t} \right)^{\frac{1}{\sigma}} \exp\left(\varepsilon_t^r\right), \tag{27}$$

$$w_t = \left(1-\alpha_0\right) \left(\frac{\Gamma_t^L Y_0^*}{\Gamma_0^L L_0} \right)^{\frac{\sigma-1}{\sigma}} \left(\frac{Y_t^*}{L_t} \right)^{\frac{1}{\sigma}} \exp\left(\varepsilon_t^w\right), \tag{28}$$

where the measurement errors are assumed to be distributed as $\varepsilon_t^r \sim N\left(0, \sigma_r^2\right)$ and $\varepsilon_t^w \sim N\left(0, \sigma_w^2\right)$. Thus, as in León-Ledesma et al. (2010a), shocks to the equilibrium real factor payments are assumed to be uncorrelated. We must notice, however, that the shocks $\varepsilon_t^{\Gamma^L}$ and $\varepsilon_t^{\Gamma^K}$ appear in both (27) and (28) (through Γ^L, Γ^K, and Y^*), which makes the correlation structure of the output and the real factor payments rather intricate.

The final step is to use the accounting identity $Y_t \equiv r_t K_t + w_t L_t$, together with the drawn realizations of K_t, L_t, r_t, and w_t, to obtain the observed output Y_t for $t = 1,\dots,T$. As discussed in León-Ledesma et al. (2010a), these steps ensure that the data are consistent with national-accounting practice, which, in turn, ensures that the shares of capital and labour sum to unity.[12]

The parameter values used to obtain the simulated series are given in Table 1. Most parameter choices are borrowed from León-Ledesma et al. (2010a). We use a variety of values for σ to learn how the local identification and endogeneity issues influence the estimation results. In particular, we start from a low 0.2 and proceed to 0.5, a value suggested by the weight of empirical evidence, surveyed by Chirinko (2008). We continue with the values 0.7, 0.8 and 0.9 to explore the behaviour

[12] In short, the experiment takes the following steps, which are repeated M times: (*i*) Generate stochastic paths for the labour, capital and technology series. (*ii*) Use these paths and equations (26)-(28) to obtain the series for the potential output and the real factor payments. (*iii*) Calculate the observed output from the accounting identity using the capital, labour and real factor payments series. (*iv*) Simulate the Markov chain of the model parameters by applying the system approach, explained in Section 2, to the simulated series. (*vi*) Calculate the point estimates of the parameters from the Markov chain and return to step *i*.

of the parameter estimates in near Cobb-Douglass production, and, finally, use the value 1.3 exceeding unity. We fix the sum of the technical progress parameters at 2% (per year) and experiment with a variety of values for γ_K and γ_L. The labour force is set to grow at 1.5% per year (l = 0.015), and, by following León-Ledesma et al. (2010a), the growth rate of capital is set at $k = l + \gamma_L$. The distribution parameter α is set at 0.4.[13]

4.2 Results

Tables 2–3 report the Monte Carlo averages and standard errors of the point estimates $\hat{\sigma}$, $\hat{\gamma}_L$, and $\hat{\gamma}_K$ obtained by the Bayesian system method, whereas Tables 4–5 provide the results under the restriction $\sigma_{12} = \sigma_{13} = \sigma_{32} = 0$. The results based on the prior standard deviation $\kappa_K = \kappa_L = 1/20$ are given in Tables 2 and 4, and the results based on $\kappa_K = \kappa_L = 1/10$ are given in Tables 3 and 5. The correlation coefficients $\rho_{13} = \sigma_{13}\big/\sqrt{\sigma_{11}\sigma_{33}}$ and $\rho_{23} = \sigma_{23}\big/\sqrt{\sigma_{22}\sigma_{33}}$, also reported in Tables 2–3, are calculated from the posterior distribution of Σ. The graphical summaries of the distributions of the point estimates are given in Figures 1–3 for the combinations $\gamma_L = 0.015$ and $\gamma_K = 0.005$, and $\gamma_L = 0.005$ and $\gamma_K = 0.015$ (based on $\kappa_K = \kappa_L = 1/20$).[14]

The estimates of ρ_{13} and ρ_{23} show that the errors of the supply side equations are strongly correlated. Thus, (5)-(7) form a system of simultaneous equations, and we know that in such a case the FGLS estimator (linear or nonlinear) is biased. Estimates of ρ_{23} are large in absolute value with low elasticity values, and they decrease as σ increases, whereas ρ_{23} tends to be large with low or high values of σ. Notice, however, that in the extreme case $\sigma = 0.2$ the Bayesian method has difficulties in obtaining accurate estimates of ρ_{13}. We remark that the system approach is in this particular case strongly misspecified due to the terms $(\sigma-1)\varepsilon_t^{\Gamma^L}\big/\sigma$ and $(\sigma-1)\varepsilon_t^{\Gamma^K}\big/\sigma$, which appear (partly non-log-linearly) on the right hand sides of (26)-(28).[15]

[13] The standard deviations of shocks are borrowed from León-Ledesma et al. (2010a, b). For capital and labour shocks the value 0.1 is used, while the standard deviations of technology shocks are set at 0.01 when $\gamma_i = 0$, and at 0.05 when $\gamma_i \neq 0$. For shocks to real wage and real interest rate (the user cost of capital) we use the standard deviations of their detrended and demeaned values, respectively, in the US economy over 1950–2000. These provide the values 0.05 and 0.1, respectively. We also performed the experiments with different plausible values of the standard deviations. The conclusions remained intact irrespective of used values.

[14] The densities of the other combinations (and those based on $\kappa_K = \kappa_L = 1/10$) are available upon request. Nevertheless, the reported cases provide a fairly general picture of the behavior of the point estimates.

[15] Note that at $\sigma = 0.2$ the "effective" standard deviations of these shocks are $4\cdot\sigma_\Gamma L$ and $4\cdot\sigma_\Gamma K$, respectively.

Given these results, it is not surprising that the exogeneity restriction $\sigma_{12} = \sigma_{13} = \sigma_{32} = 0$ corrupts the parameter estimates. In particular, Tables 3 and 5 show that $\hat{\sigma}$ is systematically biased towards unity under $\sigma_{12} = \sigma_{13} = \sigma_{32} = 0$. Smaller values of σ tend to stir up this bias, whereas in the near Cobb-Douglass case bias is relatively small. Regarding $\hat{\gamma}_L$ and $\hat{\gamma}_K$, we observe that their distributions are much flatter, having very long tails, under the exogeneity restriction.

The local identification problem is also well demonstrated in Figures 1–3 and Tables 2 and 4. Notably, the distributions of $\hat{\gamma}_L$ and $\hat{\gamma}_K$ have long fat tails in the near Cobb-Douglas case, which can clearly be seen by comparing the cases $\sigma = 0.5$ and $\sigma = 0.9$. Furthermore, the standard errors of $\hat{\gamma}_L$ and $\hat{\gamma}_K$ based on the tight prior parameterization $\kappa_K = \kappa_L = 1/20$ are considerably lower than their loose prior counterparts ($\kappa_K = \kappa_L = 1/10$), although in both cases these parameters are accurately estimated on average. These results indicate that informative priors facilitate the joint identification of γ_L and γ_K, and increase estimation accuracy, especially in the near Cobb-Douglass case. The local identification problem also tends to appear in the slightly upward biased estimates of σ. This bias becomes negligible when the true value of σ is 0.8, and it virtually disappears when σ = 0.7 (results are available on request).

Finally, when the true value of σ is in the range of 0.4–0.6, as suggested by the survey of Chirinko (2008), the Bayesian supply side method provides practically unbiased parameter estimates for all the combinations of γ_L and γ_K. The method also provides good parameter estimates for all the other combinations of σ, γ_L, and γ_K, although the estimates of σ tend to be somewhat upward biased in the extreme case $\sigma = 0.2$. As already discussed, the system approach is in this particular case strongly misspecified. The resulting increase in estimation uncertainty is revealed in the long right tail of the distribution of $\hat{\sigma}$, and in the long tails of the distributions of $\hat{\gamma}_L$ and $\hat{\gamma}_K$, given in Figures (1)-(3) (solid lines). The medians of these estimates are nevertheless close to their true values.

5. Conclusion

This paper has shown that the CES production function and the first-order conditions of profit maximizing form a (nonlinear) system of simultaneous equations, and that the information about this endogeneity should, and can, be introduced into the employed econometric model. It has further pointed out that the estimation approach of Klump et al. (2007a, 2007b, 2008) and León-Ledesma et al. (2010a) is invalid and contradictory because it fails to incorporate this knowledge into estimation.

Table 1. Parameter values for the data generating process

Parameter	Values
Substitution elasticity, σ	0.2, 0.5, 0.7, 0.8, 0.9, 1.3
Growth rate of capita-augmenting technical progress, γ_K	0.00, 0.005, 0.01, 0.015, 0.02
Growth rate of labour-augmenting technical progress, γ_L	0.02, 0.015, 0.01, 0.005, 0.00
Distribution parameter, α	0.4
Labour force growth rate, l	0.015
Capital stock growth rate, k	$l + \gamma_L$
Standard deviations of labour and capital shocks, σ_K, σ_L	0.1
Std. of capital-augmenting technical progress shock, $\sigma_{\Gamma K}$	0.01 for $\gamma_K = 0$; 0.05 for $\gamma_K \neq 0$
Std. deviation of labour-augmenting technical progress shock, $\sigma_{\Gamma L}$	0.01 for $\gamma_L = 0$; 0.05 for $\gamma_L \neq 0$
Standard deviation of real interest rate shock, σ_r	0.1
Standard deviation of real wage shock, σ_w	0.05

Notes: The number of the Monte Carlo replications is M = 1000, the sample sizes are 50 and 100, and $\gamma_K + \gamma_L = 0.02$.

Table 2. Monte Carlo results when Σ is unrestricted and the prior std. $\kappa_1 = \kappa_2 = 1/20$.

		$\sigma = 0.2$	$\sigma = 0.5$	$\sigma = 0.8$	$\sigma = 0.9$	$\sigma = 1.3$
		The averages of the point estimates (and the associated standard errors)				
	$\hat{\sigma}$	0.230 (0.040)	0.502 (0.017)	0.817 (0.046)	0.924 (0.053)	1.310 (0.126)
	$\hat{\gamma}_L$	0.018 (0.007)	0.020 (0.002)	0.020 (0.005)	0.019 (0.010)	0.020 (0.004)
γ_L =0.020	$\hat{\gamma}_K$	-0.001 (0.003)	-0.000 (0.001)	0.000 (0.007)	0.002 (0.015)	0.000 (0.006)
γ_K = 0.000	$\hat{\vartheta}$	1.056 (0.094)	1.020 (0.018)	1.005 (0.005)	1.002 (0.003)	0.994 (0.004)
	$\hat{\rho}_{13}$	-0.409 (0.651)	-0.406 (0.197)	0.156 (0.149)	0.247 (0.138)	0.420 (0.131)
	$\hat{\rho}_{23}$	-0.884 (0.240)	-0.905 (0.028)	-0.636 (0.092)	-0.509 (0.112)	-0.083 (0.173)
	$\hat{\sigma}$	0.237 (0.048)	0.502 (0.023)	0.813 (0.046)	0.924 (0.053)	1.304 (0.131)
	$\hat{\gamma}_L$	0.015 (0.003)	0.015 (0.001)	0.015 (0.004)	0.015 (0.009)	0.015 (0.004)
γ_L = 0.015	$\hat{\gamma}_K$	0.005 (0.004)	0.005 (0.002)	0.006 (0.007)	0.006 (0.013)	0.005 (0.006)
γ_K = 0.005	$\hat{\vartheta}$	1.069 (0.064)	1.020 (0.018)	1.005 (0.006)	1.002 (0.003)	0.994 (0.005)
	$\hat{\rho}_{13}$	-0.337 (0.642)	-0.446 (0.187)	0.070 (0.151)	0.211 (0.141)	0.449 (0.127)
	$\hat{\rho}_{23}$	-0.792 (0.336)	-0.861 (0.043)	-0.593 (0.105)	-0.479 (0.117)	-0.095 (0.177)
	$\hat{\sigma}$	0.238 (0.046)	0.504 (0.024)	0.814 (0.046)	0.922 (0.051)	1.313 (0.136)
	$\hat{\gamma}_L$	0.010 (0.004)	0.010 (0.001)	0.010 (0.004)	0.010 (0.008)	0.010 (0.004)
γ_L = 0.010	$\hat{\gamma}_K$	0.009 (0.003)	0.010 (0.002)	0.011 (0.006)	0.010 (0.012)	0.010 (0.006)
γ_K = 0.010	$\hat{\vartheta}$	1.076 (0.076)	1.020 (0.017)	1.006 (0.006)	1.002 (0.003)	0.994 (0.006)
	$\hat{\rho}_{13}$	-0.187 (0.664)	-0.440 (0.183)	0.080 (0.147)	0.206 (0.142)	0.462 (0.130)
	$\hat{\rho}_{23}$	-0.842 (0.295)	-0.861 (0.041)	-0.599 (0.096)	-0.479 (0.117)	-0.111 (0.175)
	$\hat{\sigma}$	0.246 (0.054)	0.506 (0.026)	0.813 (0.048)	0.922 (0.052)	1.305 (0.128)
	$\hat{\gamma}_L$	0.006 (0.003)	0.005 (0.001)	0.004 (0.004)	0.006 (0.009)	0.005 (0.005)
γ_L = 0.005	$\hat{\gamma}_K$	0.014 (0.005)	0.015 (0.002)	0.016 (0.007)	0.014 (0.014)	0.016 (0.006)
γ_K = 0.015	$\hat{\vartheta}$	1.084 (0.072)	1.023 (0.020)	1.007 (0.007)	1.003 (0.004)	0.993 (0.006)
	$\hat{\rho}_{13}$	-0.044 (0.679)	-0.398 (0.203)	0.070 (0.147)	0.199 (0.142)	0.469 (0.124)
	$\hat{\rho}_{23}$	-0.875 (0.248)	-0.864 (0.043)	-0.592 (0.100)	-0.473 (0.119)	-0.119 (0.171)
	$\hat{\sigma}$	0.254 (0.062)	0.505 (0.024)	0.812 (0.046)	0.917 (0.053)	1.310 (0.116)
	$\hat{\gamma}_L$	-0.001 (0.002)	-0.001 (0.001)	-0.001 (0.004)	0.001 (0.009)	0.000 (0.004)
γ_L = 0.000	$\hat{\gamma}_K$	0.017 (0.009)	0.020 (0.003)	0.021 (0.007)	0.019 (0.015)	0.020 (0.006)
γ_K = 0.020	$\hat{\vartheta}$	1.046 (0.067)	1.019 (0.016)	1.007 (0.008)	1.003 (0.004)	0.991 (0.008)
	$\hat{\rho}_{13}$	-0.286 (0.614)	-0.496 (0.151)	0.055 (0.150)	0.212 (0.134)	0.541 (0.119)
	$\hat{\rho}_{23}$	-0.857 (0.234)	-0.811 (0.056)	-0.563 (0.102)	-0.488 (0.109)	-0.277 (0.158)

Notes: Substitution elasticity, σ, growth rate of capita-augmenting technical progress, γ_K, growth rate of labour-augmenting technical progress, γ_L, normalization constant, □, elements of structural form covariance matrix, $\Sigma = \sigma_{ij}$, correlation coefficients, $\rho_{ij} = \sigma_{ij} \big/ \sqrt{\sigma_{ii}\sigma_{jj}}$.

Table 3. Monte Carlo results when Σ is unrestricted and the prior std. $\kappa_1 = \kappa_2 = 1/10$.

		$\sigma = 0.2$	$\sigma = 0.5$	$\sigma = 0.8$	$\sigma = 0.9$	$\sigma = 1.3$
		The averages of the point estimates (and the associated standard errors)				
$\gamma_L = 0.020$ $\gamma_K = 0.000$	$\hat{\sigma}$	0.231 (0.041)	0.502 (0.018)	0.818 (0.050)	0.930 (0.052)	1.302 (0.129)
	$\hat{\gamma}_L$	0.019 (0.006)	0.020 (0.001)	0.020 (0.006)	0.019 (0.015)	0.020 (0.005)
	$\hat{\gamma}_K$	-0.001 (0.003)	-0.000 (0.001)	0.000 (0.008)	0.002 (0.022)	0.000 (0.008)
	$\hat{\vartheta}$	1.053 (0.088)	1.016 (0.015)	1.005 (0.005)	1.002 (0.003)	0.994 (0.005)
	$\hat{\rho}_{13}$	-0.420 (0.651)	-0.411 (0.191)	0.161 (0.143)	0.250 (0.134)	0.416 (0.130)
	$\hat{\rho}_{23}$	-0.884 (0.238)	-0.904 (0.027)	-0.638 (0.098)	-0.499 (0.115)	-0.087 (0.174)
$\gamma_L = 0.015$ $\gamma_K = 0.005$	$\hat{\sigma}$	0.234 (0.043)	0.503 (0.022)	0.819 (0.052)	0.926 (0.052)	1.312 (0.139)
	$\hat{\gamma}_L$	0.015 (0.004)	0.015 (0.001)	0.015 (0.006)	0.014 (0.014)	0.015 (0.006)
	$\hat{\gamma}_K$	0.004 (0.004)	0.005 (0.002)	0.006 (0.009)	0.006 (0.021)	0.005 (0.009)
	$\hat{\vartheta}$	1.071 (0.071)	1.019 (0.017)	1.005 (0.005)	1.002 (0.003)	0.994 (0.005)
	$\hat{\rho}_{13}$	-0.333 (0.646)	-0.449 (0.185)	0.091 (0.150)	0.206 (0.139)	0.460 (0.131)
	$\hat{\rho}_{23}$	-0.783 (0.342)	-0.859 (0.042)	-0.597 (0.108)	-0.470 (0.117)	-0.107 (0.175)
$\gamma_L = 0.010$ $\gamma_K = 0.010$	$\hat{\sigma}$	0.239 (0.048)	0.504 (0.023)	0.820 (0.055)	0.930 (0.053)	1.301 (0.126)
	$\hat{\gamma}_L$	0.010 (0.003)	0.010 (0.001)	0.009 (0.007)	0.010 (0.016)	0.010 (0.005)
	$\hat{\gamma}_K$	0.009 (0.004)	0.010 (0.002)	0.011 (0.012)	0.010 (0.024)	0.010 (0.008)
	$\hat{\vartheta}$	1.074 (0.064)	1.022 (0.019)	1.006 (0.006)	1.002 (0.004)	0.994 (0.006)
	$\hat{\rho}_{13}$	-0.175 (0.656)	-0.413 (0.206)	0.077 (0.150)	0.217 (0.139)	0.458 (0.125)
	$\hat{\rho}_{23}$	-0.849 (0.264)	-0.865 (0.043)	-0.587 (0.110)	-0.473 (0.121)	-0.117 (0.167)
$\gamma_L = 0.005$ $\gamma_K = 0.015$	$\hat{\sigma}$	0.247 (0.069)	0.506 (0.023)	0.815 (0.047)	0.928 (0.051)	1.302 (0.130)
	$\hat{\gamma}_L$	0.006 (0.003)	0.005 (0.001)	0.004 (0.005)	0.004 (0.014)	0.004 (0.006)
	$\hat{\gamma}_K$	0.013 (0.008)	0.015 (0.002)	0.016 (0.008)	0.016 (0.021)	0.016 (0.009)
	$\hat{\vartheta}$	1.084 (0.074)	1.023 (0.021)	1.007 (0.007)	1.003 (0.004)	0.993 (0.006)
	$\hat{\rho}_{13}$	-0.070 (0.682)	-0.399 (0.203)	0.075 (0.148)	0.204 (0.144)	0.458 (0.129)
	$\hat{\rho}_{23}$	-0.873 (0.233)	-0.864 (0.041)	-0.592 (0.103)	-0.467 (0.116)	-0.118 (0.175)
$\gamma_L = 0.000$ $\gamma_K = 0.020$	$\hat{\sigma}$	0.258 (0.066)	0.506 (0.025)	0.815 (0.050)	0.923 (0.051)	1.313 (0.126)
	$\hat{\gamma}_L$	-0.001 (0.002)	-0.001 (0.001)	-0.001 (0.005)	-0.001 (0.014)	-0.001 (0.006)
	$\hat{\gamma}_K$	0.016 (0.010)	0.020 (0.003)	0.022 (0.009)	0.021 (0.021)	0.021 (0.008)
	$\hat{\vartheta}$	1.044 (0.068)	1.020 (0.016)	1.007 (0.007)	1.003 (0.004)	0.991 (0.008)
	$\hat{\rho}_{13}$	-0.283 (0.624)	-0.486 (0.152)	0.073 (0.144)	0.211 (0.136)	0.540 (0.109)
	$\hat{\rho}_{23}$	-0.862 (0.219)	-0.808 (0.055)	-0.567 (0.103)	-0.479 (0.114)	-0.275 (0.156)

Notes: Substitution elasticity, σ, growth rate of capita-augmenting technical progress, γ_K, growth rate of labour-augmenting technical progress, γ_L, normalization constant, □, elements of structural form covariance matrix, $\Sigma = \sigma_{ij}$, correlation coefficients, $\rho_{ij} = \sigma_{ij} \big/ \sqrt{\sigma_{ii}\sigma_{jj}}$.

Table 4. Monte Carlo results when $\sigma_{12} = \sigma_{13} = \sigma_{32} = 0$ and the prior std. $\kappa_1 = \kappa_2 = 1/20$.

		$\sigma = 0.2$	$\sigma = 0.5$	$\sigma = 0.8$	$\sigma = 0.9$	$\sigma = 1.3$
		The averages of the point estimates (and the associated standard errors)				
$\gamma_L = 0.020$ $\gamma_K = 0.000$	$\hat{\sigma}$	0.345 (0.145)	0.554 (0.036)	0.846 (0.047)	0.939 (0.045)	1.242 (0.087)
	$\hat{\gamma}_L$	0.023 (0.013)	0.020 (0.003)	0.020 (0.006)	0.019 (0.010)	0.020 (0.005)
	$\hat{\gamma}_K$	0.004 (0.018)	0.001 (0.005)	0.000 (0.009)	0.002 (0.015)	-0.000 (0.007)
	$\hat{\vartheta}$	1.047 (0.032)	1.022 (0.016)	1.006 (0.005)	1.003 (0.003)	0.994 (0.005)
$\gamma_L = 0.015$ $\gamma_K = 0.005$	$\hat{\sigma}$	0.375 (0.160)	0.569 (0.042)	0.851 (0.046)	0.941 (0.044)	1.237 (0.087)
	$\hat{\gamma}_L$	0.017 (0.013)	0.015 (0.003)	0.014 (0.006)	0.014 (0.010)	0.015 (0.005)
	$\hat{\gamma}_K$	0.013 (0.021)	0.007 (0.006)	0.007 (0.009)	0.007 (0.014)	0.005 (0.007)
	$\hat{\vartheta}$	1.046 (0.031)	1.022 (0.017)	1.007 (0.007)	1.003 (0.003)	0.994 (0.005)
$\gamma_L = 0.010$ $\gamma_K = 0.010$	$\hat{\sigma}$	0.375 (0.151)	0.571 (0.043)	0.852 (0.044)	0.938 (0.045)	1.243 (0.088)
	$\hat{\gamma}_L$	0.010 (0.009)	0.009 (0.003)	0.008 (0.006)	0.009 (0.010)	0.009 (0.005)
	$\hat{\gamma}_K$	0.020 (0.021)	0.013 (0.006)	0.013 (0.010)	0.012 (0.015)	0.011 (0.008)
	$\hat{\vartheta}$	1.05 (0.030)	1.024 (0.018)	1.008 (0.007)	1.003 (0.004)	0.993 (0.006)
$\gamma_L = 0.005$ $\gamma_K = 0.015$	$\hat{\sigma}$	0.408 (0.192)	0.571 (0.041)	0.848 (0.044)	0.940 (0.043)	1.240 (0.085)
	$\hat{\gamma}_L$	0.004 (0.008)	0.004 (0.003)	0.003 (0.006)	0.004 (0.010)	0.003 (0.005)
	$\hat{\gamma}_K$	0.029 (0.027)	0.019 (0.007)	0.019 (0.010)	0.017 (0.015)	0.017 (0.007)
	$\hat{\vartheta}$	1.042 (0.030)	1.024 (0.018)	1.009 (0.008)	1.004 (0.004)	0.993 (0.006)
$\gamma_L = 0.000$ $\gamma_K = 0.020$	$\hat{\sigma}$	0.382 (0.147)	0.562 (0.035)	0.840 (0.042)	0.932 (0.045)	1.269 (0.093)
	$\hat{\gamma}_L$	-0.001 (0.007)	-0.001 (0.002)	-0.002 (0.006)	-0.001 (0.010)	-0.001 (0.005)
	$\hat{\gamma}_K$	0.030 (0.021)	0.024 (0.005)	0.024 (0.009)	0.021 (0.016)	0.021 (0.007)
	$\hat{\vartheta}$	1.036 (0.028)	1.025 (0.016)	1.009 (0.008)	1.004 (0.004)	0.991 (0.008)

Notes: Substitution elasticity, σ, growth rate of capita-augmenting technical progress, γ_K, growth rate of labour-augmenting technical progress, γ_L, normalization constant, ϑ, elements of structural form covariance matrix, $\Sigma = \sigma_{ij}$.

Table 5. Monte Carlo results when $\sigma_{12} = \sigma_{13} = \sigma_{32} = 0$ and the prior std. $\kappa_1 = \kappa_2 = 1/10$.

		$\sigma = 0.2$	$\sigma = 0.5$	$\sigma = 0.8$	$\sigma = 0.9$	$\sigma = 1.3$
		The averages of the point estimates (and the associated standard errors)				
$\gamma_L = 0.020$ $\gamma_K = 0.000$	$\hat{\sigma}$	0.362 (0.165)	0.556 (0.038)	0.848 (0.051)	0.942 (0.048)	1.241 (0.090)
	$\hat{\gamma}_L$	0.023 (0.015)	0.020 (0.003)	0.020 (0.008)	0.019 (0.017)	0.020 (0.006)
	$\hat{\gamma}_K$	0.007 (0.026)	0.001 (0.005)	0.000 (0.012)	0.002 (0.025)	-0.001 (0.009)
	$\hat{\vartheta}$	1.044 (0.029)	1.022 (0.017)	1.006 (0.005)	1.003 (0.004)	0.994 (0.006)
$\gamma_L = 0.015$ $\gamma_K = 0.005$	$\hat{\sigma}$	0.369 (0.166)	0.572 (0.043)	0.853 (0.048)	0.943 (0.045)	1.236 (0.084)
	$\hat{\gamma}_L$	0.017 (0.013)	0.015 (0.004)	0.014 (0.008)	0.014 (0.017)	0.014 (0.007)
	$\hat{\gamma}_K$	0.014 (0.028)	0.007 (0.006)	0.007 (0.012)	0.006 (0.025)	0.006 (0.010)
	$\hat{\vartheta}$	1.047 (0.032)	1.023 (0.017)	1.007 (0.006)	1.004 (0.004)	0.994 (0.005)
$\gamma_L = 0.010$ $\gamma_K = 0.010$	$\hat{\sigma}$	0.378 (0.162)	0.571 (0.042)	0.851 (0.047)	0.942 (0.044)	1.236 (0.089)
	$\hat{\gamma}_L$	0.010 (0.011)	0.009 (0.004)	0.009 (0.008)	0.009 (0.017)	0.009 (0.007)
	$\hat{\gamma}_K$	0.021 (0.029)	0.013 (0.007)	0.013 (0.013)	0.013 (0.026)	0.011 (0.011)
	$\hat{\vartheta}$	1.045 (0.033)	1.023 (0.017)	1.007 (0.007)	1.004 (0.005)	0.993 (0.006)
$\gamma_L = 0.005$ $\gamma_K = 0.015$	$\hat{\sigma}$	0.408 (0.179)	0.575 (0.047)	0.852 (0.046)	0.940 (0.042)	1.239 (0.088)
	$\hat{\gamma}_L$	0.004 (0.010)	0.004 (0.003)	0.002 (0.009)	0.003 (0.017)	0.004 (0.006)
	$\hat{\gamma}_K$	0.031 (0.033)	0.019 (0.008)	0.020 (0.014)	0.019 (0.026)	0.016 (0.009)
	$\hat{\vartheta}$	1.041 (0.028)	1.025 (0.017)	1.009 (0.008)	1.004 (0.005)	0.992 (0.007)
$\gamma_L = 0.000$; $\gamma_K = 0.020$	$\hat{\sigma}$	0.378 (0.153)	0.562 (0.035)	0.843 (0.047)	0.937 (0.048)	1.265 (0.090)
	$\hat{\gamma}_L$	-0.001 (0.008)	-0.001 (0.002)	-0.003 (0.007)	-0.003 (0.014)	-0.001 (0.006)
	$\hat{\gamma}_K$	0.030 (0.027)	0.024 (0.005)	0.026 (0.011)	0.025 (0.022)	0.021 (0.008)
	$\hat{\vartheta}$	1.038 (0.029)	1.024 (0.016)	1.009 (0.008)	1.004 (0.005)	0.992 (0.007)

Notes: Substitution elasticity, σ, growth rate of capita-augmenting technical progress, γ_K, growth rate of labour-augmenting technical progress, γ_L, normalization constant, ϑ, elements of structural form covariance matrix, $\Sigma = \sigma_{ij}$.

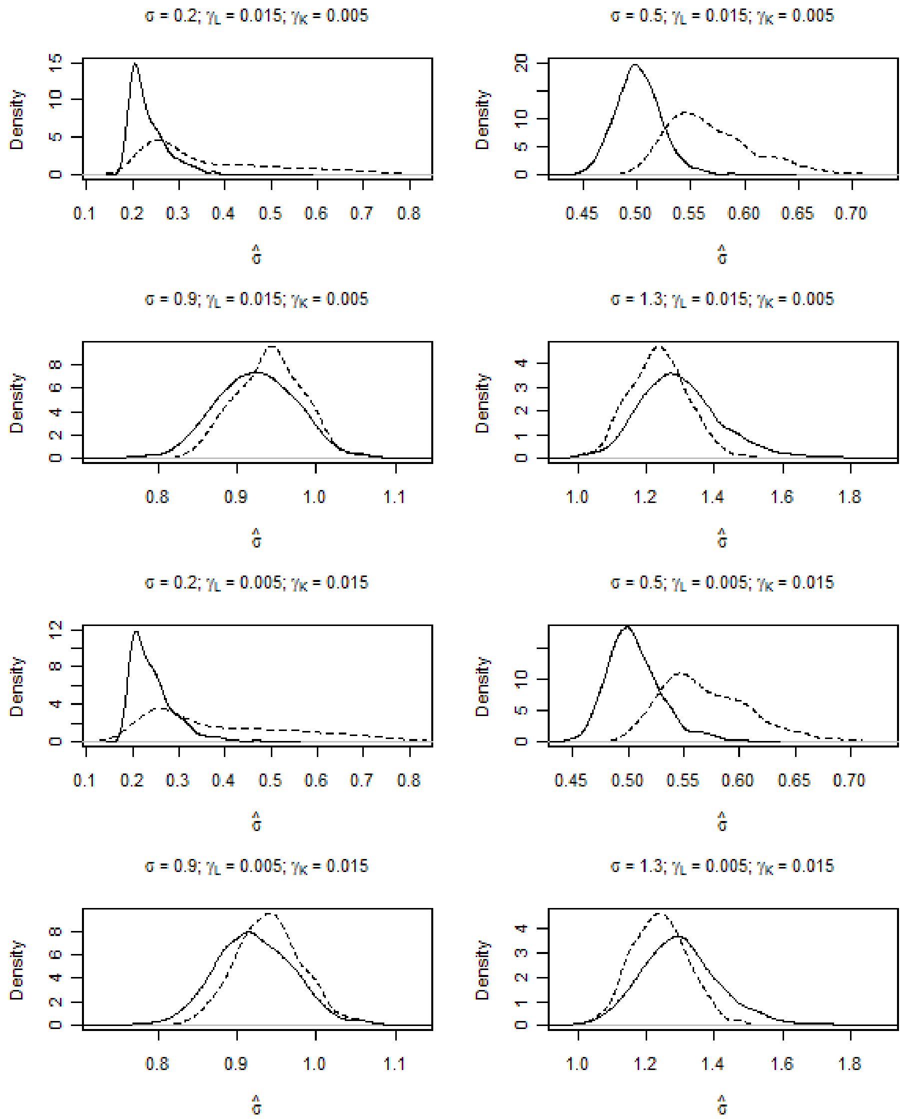

Figure 1. The distributions of the point estimates of the substitution elasticity σ (solid lines for the unrestricted structural form covariance matrix Σ; dotted lines for $\sigma_{12} = \sigma_{13} = \sigma_{32} = 0$, where σ_{ij} are the elements of Σ).

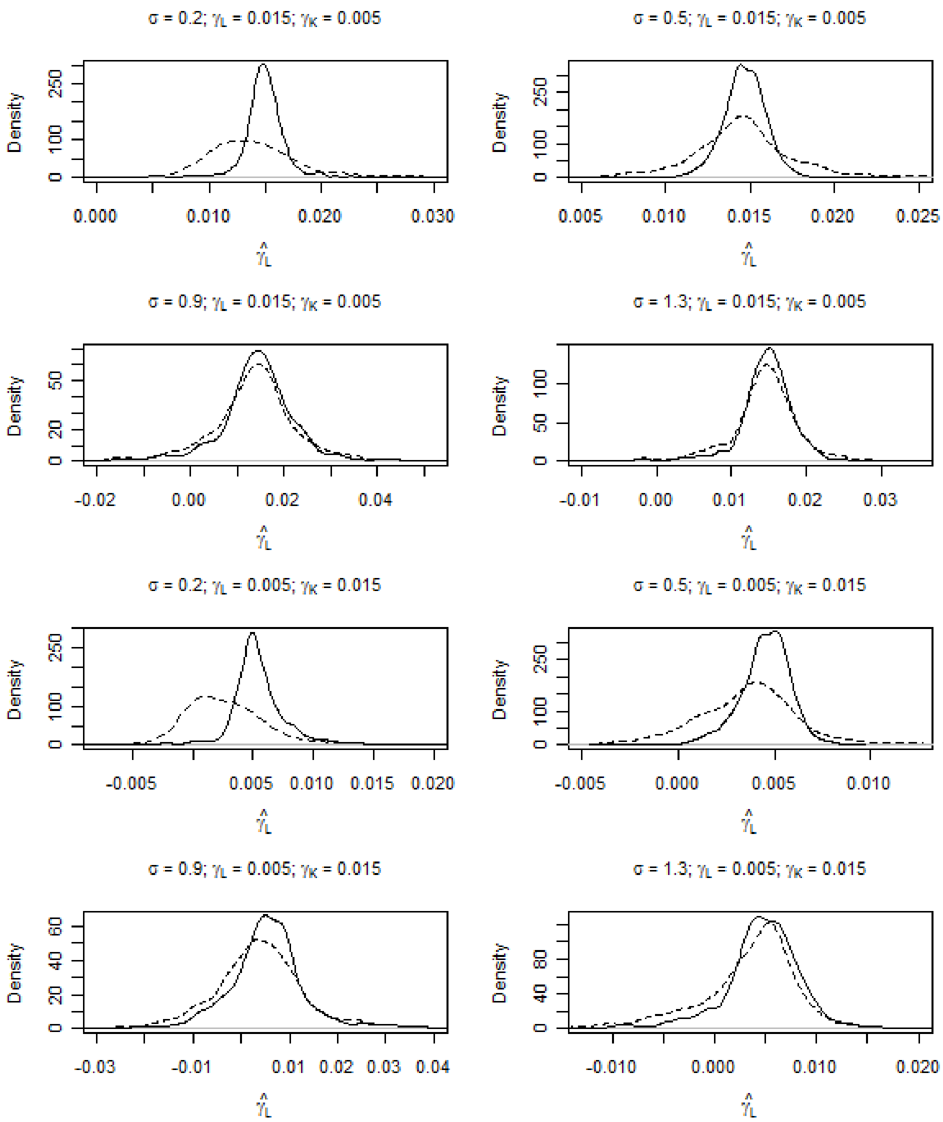

Figure 2. The distributions of the point estimates of the growth rate of labour-augmenting technical progress, γ_L, (solid lines for the unrestricted structural form covariance matrix Σ; dotted lines for $\sigma_{12} = \sigma_{13} = \sigma_{32} = 0$, where σ_{ij} are the elements of Σ).

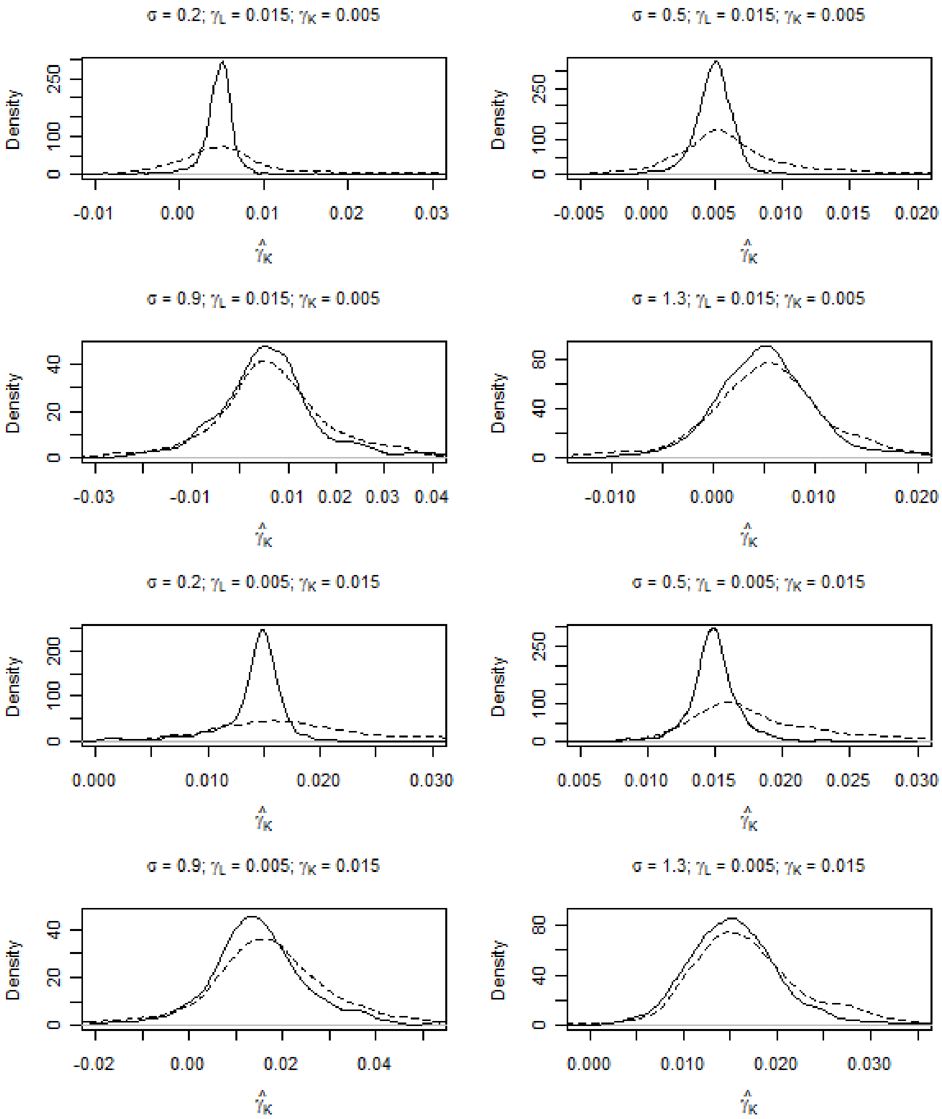

Figure 3. The distributions of the point estimates of the growth rate of capital-augmenting technical progress, γ_K, (solid lines for the unrestricted structural form covariance matrix Σ; dotted lines for $\sigma_{12} = \sigma_{13} = \sigma_{32} = 0$, where σ_{ij} are the elements of Σ).

References

Chetty, V. K. and Ulaganathan Sankar. 1969. 'Bayesian estimation of the CES production function' *The Review of Economic Studies*, 36(3) : 289–94.

Chirinko, Robert S. 2008. 'σ: The long and short of it', *Journal of Macroeconomics*, 30(2): 671–686.

de La Grandville, Oliver. 1989. 'In quest of the slutsky diamond', *American Economic Review*, 79(3): 468–481.

Kleibergen, Frank and Herman K. van Dijk. 1998. 'Bayesian Simultaneous Equations Analysis Using Reduced Rank Structures', *Econometric Theory*, 14: 701–743.

Klump, Rainer and Oliver de La Grandville. 2000. 'Economic growth and the elasticity of substitution: two theorems and some suggestions', *American Economic Review*, 90(1): 282–291.

Klump, Rainer and Harald Preissler. 2000. 'CES production functions and economic growth', *Scandinavian Journal of Economics*, 102(1): 41–56.

Klump, Rainer, Peter McAdam, and Alpo Willman. 2007a. 'Factor substitution and factor augmenting technical progress in the US: A normalized supply-side system approach', *The Review of Economics and Statistics*, 89(1): 183–192.

Klump, Rainer, Peter McAdam, and Alpo Willman. 2007b. The Long-Term SucCESs of the Neo-Classical Growth Model. *Oxford Review of Economic Policy*, 23(1): 94–114.

Klump, Rainer, Peter McAdam, and Alpo Willman. 2008. 'Unwrapping some Euro Area Growth Puzzles: Factor Substitution, Productivity and Unemployment', *Journal of Macroeconomics*, 30(2): 645–666.

Klump, Rainer, Peter McAdam, and Alpo Willman. 2011. 'The Normalized CES Production Function, Theory and Empirics', ECP Working Paper No. 1294.

Koop, Gary. 2003. *Bayesian Econometrics*, Chichester, UK: Wiley.

Koop, Gary, Rodney W. Strachan, Herman K. van Dijk, and Mattias Villani. 2005. 'Bayesian approaches to cointegratrion', Erasmus University Rotterdam, Econometric Institute, Econometric Institute Report EI 2005-13.

Lancaster, Tony. 2004. *An Introduction to Modern Bayesian Econometrics*, Oxford, UK: Blackwell.

Leon-Ledesma, Miguel A., Peter McAdam, and Alpo Willman. 2010a. 'Identifying the Elasticity of Substitution with Biased Technical Change', *American Economic Review*, 100(4): 1330–1357.

Leon-Ledesma, Miguel A., Peter McAdam, and Alpo Willman. 2010b. 'In Dubio pro CES Supply Estimation with Mis-specified Technical Change', European Central Bank, ECB Working Paper No. 1175.

McAdam, Peter and Alpo Willman. 2011. 'Technology, Utilization and Inflation: Re-assessing the New Keynesian Fundament', European Central Bank, ECB Working Paper forthcoming.

Zellner, Arnold 1971. *An introduction to Bayesian inference in econometrics*, New York: John Wiley and Sons.