Discussion Paper No. 00-70

Imposing and Testing Curvature Conditions on a Box-Cox Cost Function

Bertrand Koebel, Martin Falk, and François Laisney

Imposing and Testing Curvature Conditions on a Box-Cox Cost Function[‡]

Bertrand Koebel[*], Martin Falk[**], and François Laisney[§]

Abstract. We present a new method for imposing and testing concavity of a cost function using asymptotic least squares, which can easily be implemented even for cost functions which are nonlinear in parameters. We provide an illustration on the basis of a (generalized) Box-Cox cost function with six inputs: capital, labor disaggregated in three skill levels, energy, and intermediate materials. A parametric test of the concavity of the cost function in prices is presented, and price elasticities are compared when curvature conditions are imposed and when they are not. The results show that, although concavity is statistically rejected, the estimates are not very sensitive to its imposition. We find that substitution is stronger between the different types of labor than between any other pair of inputs.

Keywords: input demands, concavity, inequality restrictions, Box-Cox.

JEL-Classification: C33, E23, J21, J31.

[‡] We thank Herbert Buscher, Sandra Kneile and the participants in seminars in Magdeburg, Mannheim, Strasbourg and Tübingen for useful comments. Any shortcomings and errors are of course our own responsibility.

[*] Corresponding author: Otto-von-Guericke University, P.O. Box 4120, D-39016 Magdeburg. Telephone: (+49) 391-671-8431, Fax: (+49) 391-671-1218. Email: koebel@ww.uni-magdeburg.de.

[**] Centre for European Economic Research (ZEW), Mannheim, falk@zew.de.

[§] BETA-THEME, Louis Pasteur University, Strasbourg, and ZEW, Mannheim, laisney@zew.de.

1. Introduction

Most empirical studies in production analysis are based on functional forms that have to satisfy some curvature conditions in order to be compatible with microeconomic theory. The aim of this paper is to present and implement a new method for imposing price concavity of a cost function and testing this property. The main advantage of the framework is that it is easily implemented even for cost functions which are nonlinear in parameters.

Contributions in the field of production analysis often check whether concavity is fulfilled by the estimated parameters of the cost function. Since the seminal papers by Lau [1978] and by Diewert and Wales [1987], concavity is more and more often directly imposed (locally or globally) on the parameters. More recently, Ryan and Wales [1998, 2000] and Moschini [1999] discuss further techniques to impose concavity. However, few contributions test whether concavity is statistically rejected by the data. Kodde and Palm [1987] and Härdle, Hildenbrand and Jerison [1991] are notable exceptions in the context of demand analysis. Tests of the concavity assumption appear interesting from a statistical point of view, but also from an economic perspective: production units and goods considered in almost all empirical investigations are aggregates for which microeconomic properties are not necessarily valid (see for instance Koebel [2002] on this point). In this context, the a priori imposition of concavity may lead to estimation biases.

Among the alternative ways of imposing the negative semi-definiteness of a constant matrix $\overline{H}$, the one proposed by Lau [1978], and its further developments by Diewert and Wales [1987] and by Ryan and Wales [1998], are particularly attractive since they are easy to implement. Their approach consists in reparameterizing the matrix $\overline{H}$ by $H^0 = -U'U$ and estimating the parameters of the triangular matrix U instead of $\overline{H}$. The resulting matrix H^0 automatically verifies negative semi-definiteness. Several problems may arise when using this procedure. First, by applying it in turn to 29 two-digit industrial branches, Koebel [1998] found convergence problems with the nonlinear SUR estimator for the parameters of H^0 for more than half of the branches considered. These problems are even more serious when the unrestricted specification is already nonlinear in the parameters (as e.g. the Box-Cox cost function). Second, the procedure proposed by Ryan and Wales cannot be used for demand systems for which the parameters of the matrix $\overline{H}$ cannot all be identified from the reparameterization $H^0 = -U'U$ (see also Moschini [1999]). The method we outline in this paper can be applied for a very wide range of demand systems and is illustrated using a generalized Box-Cox specification which nests both the translog and the normalized quadratic functional forms.[1]

The solution we propose makes use of a minimum distance, or an asymptotic least-squares estimator, proposed by Gouriéroux, Monfort and Trognon [1985] and Kodde, Palm and Pfann [1990]. Concavity is imposed in two stages. In a first stage we estimate the unrestricted parameters to obtain estimate $\widehat{\overline{H}}$, and this will typically not be negative semi-definite. In a second stage, the difference between H^0 and $\widehat{\overline{H}}$ is minimized (for a distance measured in an appropriate metric) to obtain the concavity restricted estimates

[1] Diewert and Wales [1992] call Normalized Quadratic a functional form which was called Generalized McFadden functional form in Diewert and Wales [1987].

$\widehat{\widehat{H}}^0$. We then present a parametric test for the concavity of the cost function in prices: the method we rely on for imposing concavity can simultaneously be used for testing this assumption, by testing whether the matrix $\widehat{\widehat{H}}^0 - \widehat{\widehat{H}}$ is statistically different from zero.

These results are applied to the analysis of the impact of price, output growth and technological change on labor demand for different skill levels. Rather few studies have considered different skill classes of labor as distinct inputs in the production process. In general, labor is treated as a single aggregate input, with two kinds of undesirable consequences. First, it is only under restrictive conditions on the technology and on the evolution of prices that the different labor and material inputs can be combined into single aggregate measures. Considering aggregate labor may therefore lead to an estimation bias. Second, disaggregated information is often of interest for assessing the impact of policies to fight the high unemployment of unskilled workers (by means of wage subsidies for example). This information cannot be recovered from models considering aggregate labor. In this paper, we consider the wages of different types of labor, the prices of energy, material and capital, the level of output, and the impact of time, in order to explain the evolution of different input demands.

As concavity rejection may in fact be attributable to an inappropriate specification of the functional form, we retain a generalized Box-Cox formulation which nests several usual models. Although the generalized Box-Cox is non-linear in parameters, concavity is imposed and the parameters are estimated without great difficulties. The test for the null of concavity is computed for several specifications, and we study whether some functional forms are more likely to fulfill concavity than others. Furthermore, price elasticities are compared when curvature conditions are imposed and when they are not.

The factor demand system is estimated for 31 German manufacturing branches for the period 1978 to 1990. The skill categories are based on the highest formal qualification received: workers without any formal vocational certificate are categorized as low-skilled or unskilled; workers with a certificate from the dual vocational training system who have attained either a university level entrance degree ("Abitur") or a vocational school degree, are categorized as medium-skilled or skilled; and finally, workers with a university or technical university degree are categorized as high-skilled workers.

The observed shift in demand away from unskilled labor is widely documented in the economic literature. For Germany the situation can be visualized on Figure 1 below, which describes the evolution of the share of each skill in aggregate labor, with h_t, s_t and u_t respectively denoting high-skilled, skilled and unskilled workers and aggregate labor defined as $\ell_t = h_t + s_t + u_t$. One explanation for this shift is that technological change is skilled labor augmenting (Berman, Bound and Griliches, 1994) and that higher skilled labor is more complementary to equipment investment than lower skilled labor. Another reason for the change in employment composition is that employment changes in response to changes in wages and output vary for different skill levels (e.g., Bergström and Panas, 1992, Betts, 1997). Both effects, as well as the impact of time and of price changes, are simultaneously investigated here.

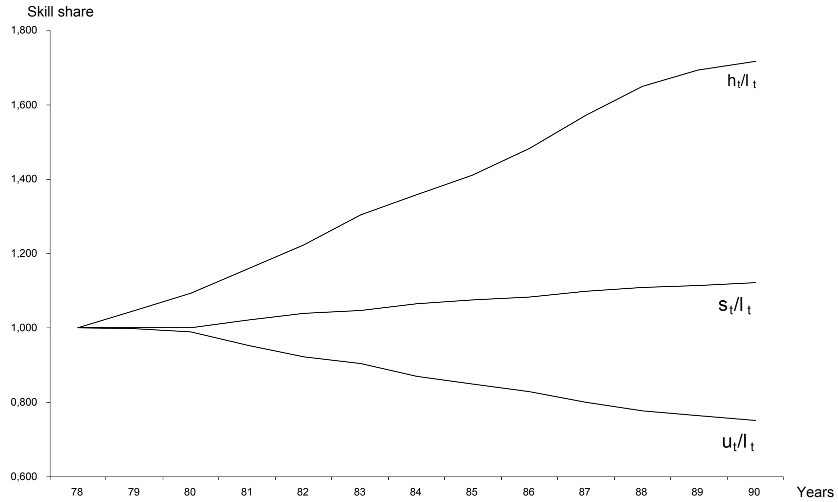

Figure 1: Evolution of the shares of three types of qualification (h=high skill, s=skill, u=unskilled) in aggregate manufacture employment (l=labor), West Germany, basis 1978=1.00.

Sections 2 and 3 are devoted to the presentation of the techniques used to impose and to test concavity, respectively. The generalized Box-Cox specification is presented in Section 4 and the results of some specification tests appear in Section 5. The results of concavity tests are examined in Section 6 and the elasticities are discussed in Section 7. Section 8 concludes.

2. Parameter estimation under concavity restriction

The technological constraint that a production unit faces is given by $f(x, z; \alpha) \leq 0$, where x is a variable input vector, z is a vector of characteristics (such as outputs and time trend) and $\alpha \in \mathcal{A} \subset \mathbb{R}^{S_\alpha}$ is the vector of unknown technological parameters. The cost function c gives the minimal value in x of the product $p'x$ which can be achieved for given prices p and the technological constraint. That is,

$$c(p, z, \alpha) = \min_x \{p'x : f(x, z, \alpha) \leq 0\}, \tag{1}$$

where $p \in \mathbb{R}_{++}^{S_p}$, $x \in \mathbb{R}_{+}^{S_x}$, $z \in \mathbb{R}_{+}^{S_z}$ and S_v denotes the dimension of a vector v. The S_x-vector of optimal input demands $x^*(p, z, \alpha)$ is obtained by applying Shephard's lemma to c. As a consequence of the rational behavior of production units, the microeconomic cost function is linearly homogeneous and concave in input prices. Concavity in prices means that the $(S_p \times S_p)$-Hessian matrix

$$H \equiv \frac{\partial^2 c(p, z; \alpha)}{\partial p \partial p'} \tag{2}$$

of the cost function will be symmetric and negative semi-definite. In addition, linear homogeneity in prices implies that

$$\frac{\partial^2 c(p,z;\alpha)}{\partial p \partial p'} p = 0, \tag{3}$$

hence, only $S_p(S_p-1)/2$ elements of H will be linearly independent.

For simplicity, linear homogeneity in prices and symmetry, which are easily imposed, will not be tested in the sequel; hence any matrix H and its estimates $\widehat{H}$ are assumed to be compatible with these properties. In general the matrix H depends on p and z; we denote $\overline{H}$ the matrix obtained from H for fixed levels of prices and characteristics: $p=\overline{p}$ and $z=\overline{z}$.

Let N denote the number of observations. The expression of the (unrestricted) model can be written as $X = X^*(\alpha)+\varepsilon$, where X^* is the $(NS_x \times 1)$ stacked vector of optimal demands x^*, X is the vector of observed inputs quantities, and ε is the $(NS_x \times 1)$ vector of added residual terms. We assume that Ω_ε, the conditional variance of ε, is consistently estimated by $\widehat{\Omega}_\varepsilon$. The (unrestricted) least squares estimator $\widehat{\alpha}$ is defined as

$$\widehat{\alpha} = \arg\min_{\alpha} (X - X^*(\alpha))' \widehat{\Omega}_\varepsilon^{-1} (X - X^*(\alpha)). \tag{4}$$

The concavity restricted least squares estimator is obtained as

$$\widehat{\alpha}^0 = \arg\min_{\alpha} \left\{ (X - X^*(\alpha))' \widehat{\Omega}_\varepsilon^{-1} (X - X^*(\alpha)) : v'\overline{H}v \le 0, \forall v \in \mathbb{R}^{S_p} \right\}. \tag{5}$$

Since the matrix H in (2) in general does not solely depend on α but also on price and output levels, concavity is only imposed locally in (5) (at $p=\overline{p}$ and $z=\overline{z}$) and will not necessarily be fulfilled at other observation points. Hence the notation $\overline{H}$ in (5). The method we propose could however be adapted to impose concavity at more than one point or globally. Several techniques for estimation under inequality constraints have recently been overviewed by Ruud [1997] and Ryan and Wales [1998]. In the sequel, we focus on a method which is attractive because of its simplicity and which can be applied to a wide range of functional specifications.

As already mentioned, symmetry and linear homogeneity are easily imposed on the cost function, and Diewert and Wales [1987] and Ryan and Wales [1998] show that restricting the parameters α to fulfill negative semi-definiteness of the matrix $\overline{H}$ is not much more difficult. Indeed, for some functional forms, the parameter vector α can be split into $\alpha = (\alpha_A', \alpha_B')'$, where α_A is a vector with $S_p(S_p-1)/2$ free parameters, whose values can be chosen to ensure negative semi-definiteness of $\overline{H}$ for any value of the remaining parameters α_B.[2] For many usual functional forms, the Hessian matrix of the cost function with respect to p can be written at a given point as $\overline{H} = \overline{A} + \overline{B}$ where the matrix $\overline{A}$ only depends on the concavity driving parameters $\alpha_A = \operatorname{vecli} \overline{A}$, and $\overline{B}$ only depends on α_B but not on α_A. The operator vecli which is introduced here stacks up a maximal subset of *linearly independent* components of a matrix. It is a slight adaptation of the operator vec, with the complication that it is not uniquely defined. However, results will not depend on the choice of the subset, provided that this choice is made once and for all, so that the

[2] For those functional forms which are flexible in the sense given by Diewert and Wales [1987], the total number of parameters S_α will always be greater than the number $S_p(S_p-1)/2$ of parameters involved in α_A, so that the decomposition $\alpha = (\alpha_A', \alpha_B')'$ is justified.

ambiguity of the definition is only superficial. We can therefore write

$$H(\overline{p}, \overline{z}; \alpha_A, \alpha_B) = A(\overline{p}, \overline{z}; \alpha_A) + B(\overline{p}, \overline{z}; \alpha_B), \tag{6}$$

or more succinctly $\overline{H} = \overline{A} + \overline{B}$. Negative semi-definiteness of $\overline{H}$ can then be obtained, for any given matrix $\overline{B}$, by choosing the free parameters α_A such that $\overline{A}$ is sufficiently negative semi-definite. For this purpose, Ryan and Wales propose to reparameterize the matrix $\overline{A}$ as $\overline{A} \equiv -U'U - \overline{B}$, where the matrix $U\,(S_p \times S_p)$ is lower triangular, and to estimate the parameters of U and $\overline{B}$ instead of $\overline{A}$ and $\overline{B}$. Negative semi-definiteness of $\overline{H}$ is then achieved by construction. As $\overline{A} = -U'U - \overline{B}$, the parameters α_A of $\overline{A}$ can be directly determined when the parameters of U and $\overline{B}$ are identified.

Let $H^0 = -U'U$ denote the restricted Hessian matrix and let $\eta_H^0(u) = \operatorname{vecli} H^0$ be the vector comprising the $S_{\eta_H} \equiv S_p(S_p - 1)/2$ free parameters of H^0. The components of η_H^0 are functions of the elements u_{ij} of U, hence the notation $\eta_H^0(u)$, with $u = \operatorname{vecli} U$. Instead of estimating the parameters $\alpha = (\alpha_A', \alpha_B')'$ of the cost function, Ryan and Wales estimate $(u', \alpha_B')'$ by solving a nonlinear least square problem of the type

$$\min_{u, \alpha_B} (X - X^*(u, \alpha_B))' \,\widehat{\Omega}_\varepsilon^{-1} (X - X^*(u, \alpha_B)). \tag{7}$$

The identification of the parameter vector $\alpha = (\alpha_A', \alpha_B')$, which is of interest for the computation of elasticities, may then be obtained from (6).

This approach presents three main drawbacks. First, convergence may be difficult to obtain. By implementing (7) in turn for 29 industrial sectors, Koebel [1998] encounters convergence problems for more than half of them. These difficulties are even more severe when the unrestricted functional form $x^*(p, z; \alpha)$ is already nonlinear in the parameters. Second, the decomposition of $\overline{H}$ as in (6) is not possible for every functional form, and the identification of the restricted parameters α_A in terms of η_H^0 and α_B is not always straightforward (see Ryan and Wales [1998] and Moschini [1999] on this last point). Third, tests for the concavity assumption are not provided.

Instead of relying on (7), we could estimate the concavity restricted parameters via the asymptotically equivalent minimum distance estimator obtained as the solution of

$$\widetilde{\alpha}^0 = \arg\min_{\alpha} \left\{ (\widehat{\alpha} - \alpha)' \,\widehat{\Omega}_\alpha^{-1} (\widehat{\alpha} - \alpha) : v'\overline{H}v \leq 0, \forall v \in \mathbb{R}^{S_p} \right\}, \tag{8}$$

where $\widehat{\alpha}$ denotes the unrestricted estimate of α, and $\widehat{\Omega}_\alpha$ is a consistent estimate of the variance matrix of $\widehat{\alpha}$. In (8), the parameters α are chosen such that the distance between the unrestricted and concavity restricted parameters is minimized. The asymptotic equivalence between the solutions of (5) and (8) is discussed by Gouriéroux and Monfort [1989, Chapter XXI].

In general, the inequality constraint $v'\overline{H}v \leq 0$ in (8) cannot be explicitly imposed on the parameters α, and the cases considered by Diewert and Wales [1987] and Ryan and Wales [1998] are exceptions rather than the rule. We therefore estimate the concavity restricted parameters in two stages, a procedure which is justified in Proposition 1. In the first stage, the parameters $\widehat{\eta}_H^0$ of the concavity restricted Hessian matrix $\widehat{\overline{H}}^0$ are determined as the solution of

$$\min_{u} \left(\widehat{\eta}_H - \eta_H^0(u)\right)' \widehat{\Omega}_H^{-1} \left(\widehat{\eta}_H - \eta_H^0(u)\right) \equiv d, \tag{9}$$

where $\widehat{\eta}_H = \text{vecli}\,\widehat{\widehat{H}}$ and $\eta_H^0(u) = \text{vecli}(-U'U)$. Let $g : \mathbb{R}^{S_\alpha} \to \mathbb{R}^{S_{\eta_H}}$ be such that $g(\alpha) \equiv \text{vecli}\,\nabla^2_{pp} c(\overline{p}, \overline{z}; \alpha)$; then a consistent estimate of the variance of $\widehat{\eta}_H$ is given by:

$$\widehat{\Omega}_H \equiv \frac{\partial g}{\partial \alpha'}(\widehat{\alpha})\,\widehat{\Omega}_\alpha \frac{\partial g'}{\partial \alpha}(\widehat{\alpha}). \tag{10}$$

Gouriéroux and Monfort [1989] and Wolak [1989] show that the minimum achieved in (8) is asymptotically equivalent to the Wald statistic d in (9).

From (9), we obtain $\widehat{\eta}_H^0$, but in most cases these estimates (in number S_{η_H}) do not allow to identify the S_α concavity restricted parameters $\widehat{\alpha}^0$ of the cost function (see footnote 2). Therefore, a second stage is needed to identify the parameters of interest $\widehat{\alpha}^0$. Identification can be achieved by adapting the Asymptotic Least Squares (ALS) framework proposed by Gouriéroux, Monfort and Trognon [1985]. The assertions and propositions which follow are proven in Appendix A. The relationship between the S_{η_H} restricted parameters $\widehat{\eta}_H^0$ and the S_α structural parameters $\widehat{\alpha}^0$ can be written as

$$\widehat{\alpha}^0 = \arg\min_\alpha \left\{ (\widehat{\alpha} - \alpha)' \,\widehat{\Omega}_\alpha^{-1} (\widehat{\alpha} - \alpha) : \widehat{\eta}_H^0 = g(\alpha) \right\}, \tag{11}$$

and the solution to this problem is asymptotically equivalent to

$$\widehat{\alpha}^0 = \widehat{\alpha} + \widehat{\Omega}_\alpha \frac{\partial g'}{\partial \alpha}(\widehat{\alpha}) \left(\frac{\partial g}{\partial \alpha'}(\widehat{\alpha})\,\widehat{\Omega}_\alpha \frac{\partial g'}{\partial \alpha}(\widehat{\alpha}) \right)^{-1} \left(\widehat{\eta}_H^0 - g(\widehat{\alpha}) \right). \tag{12}$$

From this expression it can be seen that the concavity restricted parameters $\widehat{\alpha}^0$ are equal to the unrestricted estimates $\widehat{\alpha}$ corrected by a function of the difference $\widehat{\eta}_H^0 - g(\widehat{\alpha})$ between the parameters of the concavity restricted and unrestricted Hessians. The relationships between these estimators and their properties are given in Proposition 1 below.

Proposition 1. Under the assumption that each of the problems (8), (9) and (11) has a unique solution, the following properties are verified: under the null of concavity,

(i) the solution $\widetilde{\alpha}^0$ of (8) and the solution $\widehat{\alpha}^0$ of (11) are asymptotically equivalent;

(ii) the solution $\widehat{\eta}_H^0$ of (9) is asymptotically equivalent to $g(\widetilde{\alpha}^0)$ and to $g(\widehat{\alpha}^0)$;

(iii) the minimum distances achieved in problem (8), (9) and (11) are asymptotically equivalent;

(iv) an asymptotic solution of (11) is given by (12).

Point (i) of Proposition 1 justifies our two-steps procedure for solving (8). Point (ii) allows to retain the statistic $\widehat{\eta}_H^0$ obtained by solving (9) as an estimator for $g(\alpha^0)$. Part (iii) provides a rationale for computing the LR type test for the null of concavity using the minimum value achieved in (9). Point (iv) justifies the use of (12) for the determination of the concavity restricted parameters.

Concerning the asymptotic distribution of the estimators $\widehat{\eta}_H^0$ and $\widehat{\alpha}^0$ under H_0, we must distinguish the case where the true value saturates the constraints or not. If they do not, these asymptotic distributions are $N(\alpha, \Omega_\alpha)$ and $N(\eta_H, \Omega_H)$, respectively. If some constraints are saturated by the true value, however, the distributions of $\widehat{\eta}_H^0$ and

$\widehat{\alpha}^0$ becomes quite complex; see Gouriéroux, Holly and Monfort [1982], Kodde and Palm [1986] and Wolak [1989].

A special case of this minimum distance estimator of $\widehat{\alpha}^0$ was used in Koebel [1998] for the estimation of the concavity restricted parameters of a normalized quadratic cost function: for this functional form, the matrix $B(\overline{p}, \overline{z}; \alpha_B)$ vanishes in the expression (6), and the concavity restricted parameters α_A^0 can be estimated and identified in one stage. The main advantage of using (9) and (12) rather than (7) is that convergence is obtained much more easily. As will be shown in the next section, (9) is also useful for testing the validity of the concavity restrictions.

3. Testing concavity

In consumer analysis, tests of the definiteness of a matrix are sometimes presented: for example Härdle and Hart [1992] and Härdle, Hildenbrand and Jerison [1991] present a method relying on whether the highest nonzero eigenvalue of $\widehat{\overline{H}}$ is significantly negative. Kodde and Palm [1987] prefer to consider all eigenvalues simultaneously and propose a distance test based on:

$$d_{KP} \equiv \min_{\lambda \leq 0} \left(\widehat{\lambda} - \lambda\right)' \widehat{\Sigma}^- \left(\widehat{\lambda} - \lambda\right), \tag{13}$$

where $\widehat{\lambda}$ denotes the vector of all eigenvalues of the estimated matrix $\widehat{\overline{H}}$ and $\widehat{\Sigma}$ is a consistent estimate of the variance matrix of λ given by

$$\widehat{\Sigma} = \frac{\partial \lambda}{\partial \operatorname{vec}' \overline{H}} \left(\widehat{\overline{H}}\right) \widehat{\mathbf{\Omega}}_H \frac{\partial \lambda'}{\partial \operatorname{vec} \overline{H}} \left(\widehat{\overline{H}}\right).$$

where $\widehat{\mathbf{\Omega}}_H$ is the variance matrix of $\operatorname{vec} \widehat{\overline{H}}$.[3] A generalized inverse of $\widehat{\Sigma}$ has to be considered in (13), because $\overline{H}$ is symmetric and singular, but d_{KP} is independent of the choice of generalized inverse. The following proposition gives an asymptotically equivalent expression of the test statistic d_{KP} proposed by Kodde and Palm.

Proposition 2. Under the assumption that the eigenvalues λ are differentiable with respect to η_H,

$$d_{KP} \geq d,$$
$$d_{KP} \stackrel{a}{=} d.$$

Briefly, Proposition 2 states that in small samples $d_{KP} \geq d$, but that both statistics are asymptotically equivalent. The distance d_{KP} between restricted and unrestricted eigenvalues is asymptotically equivalent to the distance d between the elements of the estimated matrix $\widehat{\overline{H}}$ and the negative semi-definite matrix $\widehat{\overline{H}}^0$. The distance d may however be more useful than d_{KP} for three reasons. First, the computation of d is somewhat simpler since we do not have to calculate the matrix of derivatives of the eigenvalues with respect to the parameters $\partial\lambda / \partial \operatorname{vec}' \overline{H}$. Second, the statistic d can be computed even when the

[3] Note that $\widehat{\mathbf{\Omega}}_H$ is different from $\widehat{\Omega}_H$ which is the variance matrix of $\operatorname{vecli} \widehat{\overline{H}}$. While the later has full rank, the former is singular.

eigenvalues are multiple and not differentiable. Third, in the case where α can be split into $(\eta_H', \alpha_B')'$, we can directly obtain the restricted parameters $\widehat{\alpha}_A^0$ by solving (9); this is not the case when using (13).

Does the assumption that the eigenvalues are differentiable strongly restrict the applicability of Proposition 2? The following result shows that the set of matrices which have multiple eigenvalues is of measure zero, and therefore Proposition 2 can almost always be applied.

Proposition 3. (i) The eigenvalue λ_1 is differentiable with respect to η_H if and only if λ_1 is simple.

(ii) The set of all matrices $\overline{H}$ which have multiple eigenvalues is of Lebesgue measure zero.

Proposition 3 means that almost all matrices have differentiable eigenvalues. Hence, the technical problem related to the non-differentiability of some eigenvalues which may arise when Proposition 2 is applied occurs only for a small set of matrices with measure zero. The following example illustrates the proposition. Let $S_p = 2$, then a symmetric matrix $\overline{H}$ satisfying (1) and (3) at $\overline{p} = (1,1)$ is of the form

$$\begin{pmatrix} a & -a \\ -a & a \end{pmatrix},$$

and has multiple eigenvalues if and only if $a = 0$. When the distribution of the random parameter a is continuous and smooth, the occurrence of multiple eigenvalues is an unlikely event.

Gouriéroux, Holly and Monfort [1982], Kodde and Palm [1986] and Wolak [1989] have shown that under the null hypothesis, the statistic d will asymptotically follow a mixture of Chi-squared distributions:

$$\Pr\left[d \geq \underline{d}\right] \overset{a}{=} \sum_{j=0}^{S_p-1} \Pr\left[\chi^2(j) \geq \underline{d}\right] w\left(S_p - 1, S_p - 1 - j, \Omega_H\right),$$

where the weight w_j denotes the probability that j of the $S_p - 1$ eigenvalues of $\overline{H}$ are negative. As the computation of the weights in the expression of d is not straightforward, the lower and upper bounds to the critical value computed by Kodde and Palm [1986] will be used for hypothesis tests.

4. A generalized Box-Cox cost function

In order to avoid the imposition of a priori restrictions on an unknown technological structure, many researchers have relied on flexible functional forms which can be interpreted as a (local) second order approximation to an arbitrary cost function. Translog, generalized Leontief and normalized quadratic cost functions have often been used for estimating price elasticities. We consider a generalized Box-Cox cost function which nests the former usual specifications. In contrast to the Box-Cox formulations of Berndt and Khaled [1979] and of Lansink and Thijssen [1998], our specification nests both the normalized quadratic and the translog cost functions.

In order to specify our model, we apply the Box-Cox transformation to the explanatory variables p_{it} and z_{it}: for $\gamma_1 \neq 0$, let

$$\begin{aligned} Z_{jit} &= \frac{z_{jit}^{\gamma_1} - 1}{\gamma_1}, \quad j = 1, \ldots, S_z, \\ P_{jit} &= \frac{(p_{jit}/\theta_i' p_{it})^{\gamma_1} - 1}{\gamma_1}, \quad j = 1, \ldots, S_p. \end{aligned}$$

For $\gamma_1 = 0$, let $Z_{jit} = \ln z_{jit}$ and $P_{jit} = \ln (p_{jit}/\theta_i' p_{it})$. The term $\theta_i' p_{it}$ appearing in the expression of P_{jit} is introduced to guarantee that the cost function is linearly homogeneous in prices. The vector θ_i of size $S_p \times 1$ is chosen to be equal to x_{i1}/c_{i1} so that $\theta_i' p_{it}$ corresponds to a Laspeyres price index for total costs, normalized to '1' for $t = 1$. The choice of the Laspeyres cost index for normalization is appealing because the Laspeyres index is independent of the units of measurement of prices and of quantities (it satisfies the index theoretical dimensionality and commensurability axioms).

The specification of the cost function is

$$c^* (p_{it}, z_{it}; \theta_i, \alpha_i) = p_{it}' x_{i1} \left(\gamma_2 C^* (p_{it}, z_{it}; \theta_i, \beta_i) + 1\right)^{1/\gamma_2}, \tag{14}$$

for $\gamma_2 \neq 0$ and $c^* = p_{it}' x_{i1} \exp (C^*)$ for $\gamma_2 = 0$, where

$$\begin{aligned} C^* (p_{it}, z_{it}; \theta_i, \beta_i, \gamma_1) &= C (P_{it}, Z_{it}; \beta_i) \\ &= \beta_{0i} + (P_{it}', Z_{it}') B_1 + \frac{1}{2} (P_{it}', Z_{it}') B_2 \begin{pmatrix} P_{it} \\ Z_{it} \end{pmatrix} \\ &= \beta_{0i} + (P_{it}', Z_{it}') \begin{pmatrix} B_{pi} \\ B_z \end{pmatrix} + \frac{1}{2} (P_{it}', Z_{it}') \begin{pmatrix} B_{pp} & B_{pz} \\ B_{zp} & B_{zz} \end{pmatrix} \begin{pmatrix} P_{it} \\ Z_{it} \end{pmatrix}. \end{aligned} \tag{15}$$

In c^*, the technological parameters to be estimated are gathered in the vector $\alpha_i = (\beta_i', \gamma_1, \gamma_2)'$. The matrices B_1 and B_2 contain the parameters of β_i and are of size $(S_p + S_z) \times 1$, and $(S_p + S_z) \times (S_p + S_z)$, respectively. It can be directly seen that the cost function c^* is linearly homogeneous in prices. The term $p_{it}' x_{i1}$ appearing in the expression of c^* ensures both price homogeneity of degree one of the cost function and scale invariance of the estimated parameters' t-values. The sensitivity of the t-values with respect to an arbitrary scaling of the dependent variable is a problem often arising with nonlinear models (see Wooldridge [1992] for a discussion in the context of Box-Cox regression models). To understand why scale invariance holds here, consider the regression $c_{it} = c^* + \nu_{cit}$, where c denotes observed costs, c^* is defined in (14) and ν_{cit} is the realization of a random variable: changing the scaling of $c_{it} = p_{it}' x_{it}$ will similarly change the scaling of the multiplicative term $p_{it}' x_{i1}$ in the expression c^* and leave all parameter estimates unaffected.

A (locally) flexible function must be able to approximate the level, the $S_p + S_z$ first order derivatives and the $(S_p + S_z)^2$ second order derivatives of an arbitrary function at a given point. This corresponds to the number of parameters entailed by the specification (15), which thereby satisfies a necessary requirement for being flexible. Yet without further restrictions on the parameters β_i, the function C^* is not parsimoniously parameterized. Symmetry in (p_{it}, z_{it}) and homogeneity of degree zero in p_{it} imply respectively $(S_p + S_z)(S_p + S_z - 1)/2$ and $1 + S_p + S_z$ additional restrictions on C^*. Hence in order for it to be a flexible function it is only necessary that C^* entails at least $(S_p + S_z)(S_p + S_z + 1)/2$ free parameters. These additional restrictions are imposed on the parameters β_i as follows:

$$\begin{aligned} B_{pp} &= B'_{pp}, \quad B_{zz} = B'_{zz}, \quad B_{pz} = B'_{zp}, \\ \iota'_{S_p} B_{pi} &= 1, \quad \iota'_{S_p} B_{pp} = 0, \quad \iota'_{S_p} B_{pz} = 0, \end{aligned} \tag{16}$$

where ι_{S_p} denotes a $(S_p \times 1)$-vector of ones.

From (14), (15) and (16), it can be seen that several known functional forms are obtained for particular values of the parameters (γ_1, γ_2). A complete justification of the following assertions can be found in Appendix B. For $\gamma_1 = \gamma_2 = 1$, the normalized quadratic cost function (NQ) is obtained. The generalized Leontief (GL) corresponds to $\gamma_1 = 1/2$ and $\gamma_2 = 1$. The generalized square root (GSR) is obtained for $\gamma_1 = 1$ and $\gamma_2 = 2$. When $\gamma_1 \to \gamma_2 \to 0$, the translog (TL) is the limiting case. A log-linear (resp. lin-log) specification is obtained as $\gamma_1 = 1$ and $\gamma_2 \to 0$ ($\gamma_1 \to 0$ and $\gamma_2 = 1$). It is easy to see that the above Box-Cox (BC) specification is a flexible functional form: as (14) entails several flexible functional forms as special cases, the Box-Cox cost function itself is flexible.

The system of input demands $x^*(p_{it}, z_{it}; \theta_i, \alpha_i)$ is obtained through Shephard's lemma. Note that the dependence of the Laspeyres index on current prices must be taken into account in the derivation to obtain:

$$x^*_{it} = x_{i1} c^*_{it} / (p'_{it} x_{i1}) + p'_{it} x_{i1} (c^*_{it} / (p'_{it} x_{i1}))^{(1-\gamma_2)} \frac{\partial P'_{it}}{\partial p_{it}} \frac{\partial C_{it}}{\partial P_{it}}, \tag{17}$$

where

$$\frac{\partial C_{it}}{\partial P_{it}} = B_{pi} + B_{pp} P_{it} + B_{pz} Z_{it},$$

and

$$\frac{\partial P_{it}}{\partial p'_{it}} = \frac{\widehat{p}^{\gamma_1 - 1}}{(\theta'_i p_{it})^{\gamma_1}} \left(I_{S_p} - \frac{1}{\theta'_i p_{it}} p_{it} \theta'_i \right).$$

By convention $\widehat{p} \equiv \operatorname{diag}(p_{it})$ is a diagonal matrix with elements p_{ijt} on the main diagonal. We verify that

$$\frac{\partial P_{it}}{\partial p'_{it}} p_{it} = 0,$$

as a consequence of P_{it} being homogeneous of degree zero in p_{it}. Hence, for the specification (17), the adding-up condition $p'_{it} x^* = c^*$ is automatically satisfied.

The Hessian of the cost function with respect to prices is given by

$$\begin{aligned} \frac{\partial^2 c^*}{\partial p_{it} \partial p'_{it}} &= x_{i1} (c^*_{it}/p'_{it} x_{i1})^{(1-\gamma_2)} \frac{\partial C_{it}}{\partial P'_{it}} \frac{\partial P_{it}}{\partial p'_{it}} + (c^*_{it}/p'_{it} x_{i1})^{(1-\gamma_2)} \frac{\partial P'_{it}}{\partial p_{it}} \frac{\partial C_{it}}{\partial P_{it}} x'_{i1} \\ &\quad + p'_{it} x_{i1} (1-\gamma_2) (c^*_{it}/p'_{it} x_{i1})^{(1-2\gamma_2)} \left(\frac{\partial P'_{it}}{\partial p_{it}} \frac{\partial C_{it}}{\partial P_{it}} \right) \left(\frac{\partial C_{it}}{\partial P'_{it}} \frac{\partial P_{it}}{\partial p'_{it}} \right) \\ &\quad + p'_{it} x_{i1} (c^*_{it}/\theta'_i p_{it})^{(1-\gamma_2)} \left(\frac{\partial P'_{it}}{\partial p_{it}} B_{pp} \frac{\partial P_{it}}{\partial p'_{it}} + \left(\frac{\partial C_{it}}{\partial P'_{it}} \frac{\partial^2 P_{it}}{\partial p_{jt} \partial p_{ht}} \right) \right). \end{aligned} \tag{18}$$

After an evaluation at $(\overline{p}_{it}, \overline{z}_{it})$, we see that this matrix does not admit an additively separable representation such as (6). For this reason, it is more convenient to apply the method presented in Section 2 for the determination of the concavity constrained estimates. Evaluating (18) at the unrestricted parameter values $\widehat{\alpha}_i$ yields $\widehat{\eta}_H$ which can in turn be replaced in problem (9) in order to derive the minimum distance estimates $\widehat{\eta}_H^0$ for the parameters of the restricted Hessian matrix $\nabla^2_{pp} c\left(\overline{p}, \overline{z}; \widehat{\alpha}^0\right)$. Knowing the value of $\widehat{\eta}_H^0$, the restricted parameters $\widehat{\alpha}^0$ which are of interest for the computation of the different elasticities can be computed using (12).

5. Empirical implementation

We first shortly describe the data set we use and then present some preliminary results aimed to precise the sample split and the specification we rely on for testing concavity.

5.1 Data description

Given the data available, we define the vector of inputs as $x_{it} = (k_{it}, h_{it}, s_{it}, u_{it}, e_{it}, m_{it})'$ and the prices as $p_{it} = (p_{kit}, p_{hit}, p_{sit}, p_{uit}, p_{eit}, p_{mit})'$, where the labor input h_{it} denotes high-skill labor, s_{it} denotes skilled labor and u_{it} low-skilled or unskilled workers. Labor is measured in total workers (full-time equivalent). In addition, e_{it} denotes energy, m_{it} material and k_{it} capital. The subscripts t and i denote time and branch, respectively. Other explanatory variables entering the cost function are the level of production y_{it}, and a time trend t.[4] These variables are regrouped in a vector $z_{it} = (y_{it}, t)'$. The total costs of production are defined by $c_{it} = p'_{it} x_{it}$.

Within our model specification, the net capital stock is assumed to be variable. In fact there exist several economic reasons (as time to build and adjustment costs) in favor of treating capital and perhaps high skill labor as fixed or quasi-fixed inputs. Several difficulties are however related to these approaches: Gagné and Ouellette [1997] show on the basis of a Monte-Carlo analysis that meaningful estimates of the shadow value of capital may be difficult to obtain.

The user costs of capital are computed using the investment price $p_{\triangle kit}$, the nominal interest rate r_t and the depreciation rate δ_{it}:

$$p_{kit} = (r_t + \delta_{it})\, p_{\triangle kit}.$$

The depreciation rate is calculated as $\delta_{it} = 1 - (k_{it} - \Delta k_{it})/k_{i,t-1}$ where Δk_{it} denotes gross investment at constant prices. Annual interest rates are drawn from the Deutsche Bundesbank (long-term interest rate for public sector bonds). Our results differ somewhat from those obtained with the alternative formula $p_{kit} = (1 + r_t)\, p_{\triangle kit} - (1 - \delta_{it})\, p_{\triangle ki,t+1}$ for the user costs of capital.[5]

The data used consists of a panel of 31 out of 32 German two-digit manufacturing branches observed over period 1978-90. One branch (Petroleum processing, No 15) has been dropped from our sample because of the importance of taxes included in the output and the unreliability of the data available on the different skills. The choice of the period is related to the fact that energy expenditures and quantities, which are based on input-output tables, are only available from 1978 onwards. Because of data-related difficulties

4 The Box-Cox transformation is not applied to t. Hence $Z_{it} = \left(\left(y_{it}^{\gamma_1} - 1\right)/\gamma_1, t\right)$ in (15).

5 Janz and Koebel [2000] present a recent comparison of the empirical performance of alternative models of capital formation.

appearing with the German reunification, we prefer not to use post-reunification data. Most of our data were drawn from the German National Accounts.

We disaggregate the total number of employees and total labor cost into three categories by using detailed information on earnings and qualifications. Information on employment by education is taken from the Employment Register of the Federal Labor Office (Bundesanstalt für Arbeit). It contains information on employment by skill category and by branch as of June 30, for all employees paying social security contributions (each year between 1975 and 1996 is covered). Labor is divided into three groups: group 1 (high skilled) is defined as workers with a university or polytechnical degree, group 2 (skilled) is made up of those having completed vocational training as well as of technicians and foremen and the remaining group 3 (unskilled) comprises workers without formal qualifications. From this dataset we calculate the shares of the 3 skill groups in employment and multiply these proportions with total employment available for each branch from the national accounts to obtain h_{it}, s_{it} and u_{it}.

Information on earnings is taken from the IAB_S for skilled and unskilled labor and from the Federal Statistical Office (Löhne und Gehälter Statistik) for high-skilled labor.[6] From this dataset we calculate relative wages $(p_{hit}/p_{uit})_{IAB_S}$ and $(p_{sit}/p_{uit})_{IAB_S}$ and combine these figures with the national account data to obtain p_{hit}, p_{sit} and p_{uit} as

$$\begin{aligned} p_{uit} &= \frac{p_{\ell it}\ell_{it}}{\left(\frac{p_{hit}}{p_{uit}}\right)_{IAB_S} h_{it} + \left(\frac{p_{sit}}{p_{uit}}\right)_{IAB_S} s_{it} + u_{it}}, \qquad (19) \\ p_{sit} &= \left(\frac{p_{sit}}{p_{uit}}\right)_{IAB_S} p_{uit}, \\ p_{hit} &= \left(\frac{p_{hit}}{p_{uit}}\right)_{IAB_S} p_{uit}. \end{aligned}$$

This method relies on the assumption of proportionality between the wages in the IAB_S and in the national accounts, and it ensures that the wage bill $p_{hit}h_{it} + p_{sit}s_{it} + p_{uit}u_{it}$ coincides with that of the national accounts given by $p_{\ell it}\ell_{it}$.

For rendering the homoscedasticity of the vector of added residuals, ν, more plausible, the system of input demands is divided by the output levels:

$$x_{it}/y_{it} = x^* \left(p_{it}, y_{it}, t; \theta_i, \alpha_i\right)/y_{it} + \nu_{it}. \qquad (20)$$

As the parameter vector $\theta_i \equiv x_{i1}/c_{i1}$ is included as explanatory variable, the residual term ν_{it} is correlated with the θ_i for $t = 1$. In order to avoid this endogeneity problem, we drop the observations for which $t = 1$ from the regression. For the 1979-1990 period, the factor demand equations for capital, energy, material and the three types of labor are estimated with the iterative nonlinear *SUR* estimator, assuming that vector ν_{it} has zero mean and a constant variance matrix Ω_ν, and that it is uncorrelated with the regressors.[7] We thus obtain maximum likelihood estimates under the assumption $\nu \sim N(0, \Omega_\nu)$.

[6] The IAB_S dataset is a 1% random sample of all persons covered by the social security system. Depending on the year, it includes between 66,995 and 74,708 individuals working in manufacturing industries. The earnings for high-skilled workers are unfortunately top-coded in the IAB_S.

[7] Note that the matrix Ω_ν is not singular.

5.2 Preliminary results

First, the parameters α_i have been estimated by assuming that the relation (20) is valid for all branches in our sample. For $S_p = 6$, the specification (17) entails 218 free parameters (among which 186 branch dummies), which have to be estimated on the basis of $31 \times 12 \times 6 = 2232$ observations. To account for sectoral differences, the Box-Cox specification (15) includes some branch dummies β_{0i}, but in fact the remaining coefficients might also differ across branches. Given the relatively short time period available, the parameters of the Box-Cox model cannot be estimated for each branch separately. Therefore, we investigate parameter heterogeneity by estimating model (20) for different subgroups of branches. These groups are formed on the basis of similarities (a) in their production, (b) in their size, (c) in their skill structure of labor, (d) on whether they are labor or (e) capital intensive. Within each group it is assumed that technologies only differ through β_{0i} and B_{pi}, whereas across groups, technologies can differ in any of the parameters α_i.

The first sample split distinguishes three groups of branches according to the main type of production: those mainly producing (i) intermediate inputs (ii) investment goods and (iii) consumption goods. This classification is retained by the German Federal Statistical Office for the calculation of aggregate values for one-digit industries. In the second sample split, we classify the branches in 3 groups according to their level of production (at 1985 values). In the three last sample splits, we categorize the branches according to the size of some cost shares. In each case, we split the 32 branches into three groups, each entailing 10 or 11 branches, according to whether they are located in the lower, the median or upper 33.3-quantile of the distribution of a relevant variable, which is y_{it}, $p_{hit}h_{it}/c_{it}$, $(p_{hit}h_{it} + p_{sit}s_{it} + p_{uit}u_{it})/c_{it}$ and $p_{kit}k_{it}/c_{it}$ for the sample split based on output level, skill, labor and capital intensity, respectively.[8] For the groups with 10 (respectively 11) branches, there are 92 (98) parameters (among which 60 (66) are branch dummies) which have to be determined using $10 \times 12 \times 6 = 720$ (792) observations. The results of the tests for the null of identical Box-Cox technologies across branches are summarized in Table 1. In all cases, the pooled model is rejected by the likelihood ratio test. The sample splits (a) and (c) yield the highest log-likelihood values.

Even in the case where statistical tests reject the equality of some parameters, there exist arguments in favor of pooling the data. Baltagi [1995, chapter 4] recommends the use of a mean squared error (MSE) criterion for assessing the poolability of the data, rather than tests on the equality of parameters. In modelling cigarette demand, Baltagi, Griffin and Xiong [2000] found that pooled data may provide more reliable forecasts, because the "efficiency gains from pooling appear to more than offset the biases due to heterogeneity". As a long time period is necessary for such comparisons, we cannot pursue along these lines. Instead, we consider below the pooled model along with the two disaggregate models for which the highest likelihood was reached. Our choice is justified by the fact that a comparison of pooled versus disaggregate estimates may be interesting with regard to the tests for functional forms and for concavity.

[8] Since we consider intermediate material inputs as a factor of production, the industries which are not labor intensive will not automatically be capital intensive.

Table 1: Box-Cox estimates for different sample splits[1]

Split (a)	consumer goods	investment goods	intermediate goods	LR-test[2]
log-L	3651.43	2370.14	2482.76	1814.0
Split (b)	small branches	medium branches	large branches	LR-test[2]
log-L	2621.26	2674.48	2755.66	908.1
Split (c)	low skill intensive	skill intensive	high skill intensive	LR-test[2]
log-L	3245.45	2743.28	2455.91	1694.6
Split (d)	not labor intensive	labor intensive	highly labor intensive	LR-test[2]
log-L	2904.73	2658.75	2625.56	1183.4
Split (e)	not capital intensive	capital intensive	highly capital intensive	LR-test[2]
log-L	3040.00	2602.68	2578.73	1248.1

(1) See Tables C1 and C2 in Appendix C for the denomination of the industries corresponding to the different sample splits.

(2) The LR-test is calculated as $2\left(\sum_{j=1}^{3}\log \ell_j - 7597.34\right)$ where 7597.34 is the log-likelihood obtained on the pooled sample and $\log \ell_j$ denotes the log-likelihood obtained for the j^{th} group of the corresponding sample split. Under the null, this test statistic is chi-squared with 106 degrees of freedom (there are $3 \times 32 - 32 = 64$ slope parameters and $3 \times 21 - 21 = 42$ terms of the covariance matrix which may be different in the split regressions). The corresponding 5% threshold critical value is 131.03.

5.3 Tests of nested specifications

Several usual specifications of input demands are nested within the Box-Cox model and can therefore easily be tested against it. In Table 2, we provide the estimates obtained for $\widehat{\gamma}_1$ and $\widehat{\gamma}_2$, their t-statistics and the log-likelihood values obtained for the pooled and two disaggregate models (entailing each three subsamples). The upper part of Table 2 gives the result for the Box-Cox specification. Although they differ between subsamples, $\widehat{\gamma}_1$ and $\widehat{\gamma}_2$ are in all cases comprised between 0 and 1. These estimates are however statistically different from 0 and 1, which already suggests that the estimated Box-Cox is actually different from common functional forms. In two cases (subsamples (ii) in split (a) and (i) in split (c)), the assumption $\gamma_1 = \gamma_2$ cannot be rejected, which corresponds to a Box-Cox form similar to the one proposed by Berndt and Khaled [1979].

The log-likelihood values for alternative functional forms nested within the Box-Cox are reported in Table 2. Among the different usual specifications, the translog achieves the highest likelihood, followed by the normalized quadratic. However, on the basis of a likelihood-ratio test, the null hypothesis that the alternative functional form describes the technology as well as the Box-Cox is rejected for all specifications and samples considered.

As all alternative specifications are rejected, only the Box-Cox should be retained in the sequel. For assessing the disparities between alternative functional forms, both in terms of their elasticities and in their ability to satisfy concavity, we continue to consider the normalized quadratic and the translog. This permits to assess whether concavity violation is due to the choice of a particular functional form or is rejected for the bulk of the specifications.

Table 2: Log-Likelihood values of alternative specifications

Specification	Sample		γ_1 (t-value)		γ_2 (t-value)		Log-likelihood
Box-Cox	pooled		0.360	(14.1)	0.260	(14.6)	7597.34
	split (a)	(i)	0.595	(16.2)	0.347	(19.3)	8504.33
		(ii)	0.018	(0.4)	0.047	(2.3)	
		(iii)	0.541	(11.3)	0.191	(5.4)	
	split (c)	(i)	0.189	(3.1)	0.191	(8.8)	8444.64
		(ii)	0.571	(12.5)	0.146	(5.4)	
		(iii)	0.527	(15.9)	0.385	(12.0)	
Normalized Quadratic	pooled		1		1		7392.65
	split (a)						8256.20
	split (c)						8204.02
Generalized Leontief	pooled		1/2		1		7285.95
	split (a)						8126.22
	split (c)						8074.30
Generalized Square-Root	pooled		1		2		6997.64
	split (a)						7886.53
	split (c)						7860.66
Translog	pooled		$\to 0$		$\to 0$		7453.24
	split (a)						8313.98
	split (c)						8292.78
Lin-Log	pooled		1		$\to 0$		7059.10
	split (a)						7899.54
	split (c)						7952.71
Log-Lin	pooled		$\to 0$		1		7122.83
	split (a)						8135.50
	split (c)						8083.25

6. Concavity tests

In this section we first present further estimates based on the concavity unrestricted model discussed above. Then, using the method outlined in Section 2, we determine the concavity restricted parameters $\widehat{\alpha}^0$ and the test statistic $\widehat{d}$ for the null of concavity. The results are presented in Table 3. Columns three to five refer to the concavity unrestricted estimates. We evaluate the unrestricted Hessian $\nabla^2_{pp}c\,(p_{it}, z_{it}; \alpha_i)$, which is different at each observation point, and calculate the mean (over i and t) of the highest eigenvalue as well as the mean number of positive eigenvalues. The corresponding sample standard deviation is reported in parentheses. The percentage of observations violating concavity appears in column 5.

The main conclusion that can be drawn from these unrestricted estimates is that concavity is not often verified. It seems difficult to comprehensively summarize the comparisons between pooled and disaggregate estimates or between the different functional forms. Only in the cases of the BC and the TL is concavity not violated on the whole sample. With the BC, which is more flexible than the TL, it can however not be said that concavity is more often observed than with the TL. For the NQ, concavity was violated globally for any sample considered. At the more disaggregate level, however, we found a greater number of positive eigenvalues than in the pooled sample. For the BC and the TL, concavity is only observed in a very limited number of cases. It seems that concavity is a little less often violated at the disaggregate level than for the pooled sample.

Table 3: Tests of price concavity of the cost function(1)

			unrestricted estimates					restricted estimates			
			Highest		No. of positive		% of	Concavity		% of concav.	
Specif.	Sample		eigenvalue		eigenvalues		failures	test $\widehat{d}$		reject	accept
Box-Cox	pooled		0.6	(0.9)	0.8	(0.4)	100.0	36.8	(23.2)	79.0	7.5
		(i)	0.3	(0.4)	0.9	(0.3)	92.3	21.7	(12.0)	76.9	15.4
	split (a)	(ii)	5.9	(6.1)	1.9	(0.3)	100.0	20.9	(5.4)	100.0	0.0
		(iii)	2.0	(2.4)	2.7	(0.5)	100.0	16.9	(7.9)	68.5	0.0
		(i)	0.5	(1.1)	0.9	(0.7)	70.5	9.6	(13.9)	26.8	49.2
	split (c)	(ii)	2.4	(2.3)	1.9	(0.3)	100.0	37.8	(18.8)	100.0	0.0
		(iii)	1.4	(1.5)	2.4	(0.8)	100.0	16.7	(5.4)	87.5	0.0
Normalized	pooled		0.1	(0.0)	1.0	(0.0)	100.0	9.7	(1.3)	1.3	0.0
Quadratic(2)		(i)	0.3	(0.1)	1.0	(0.0)	100.0	1.9	(0.1)	0.0	100.0
	split (a)	(ii)	3.2	(0.6)	2.0	(0.0)	100.0	29.1	(0.3)	100.0	0.0
		(iii)	5.3	(0.7)	3.0	(0.0)	100.0	25.8	(0.3)	100.0	0.0
		(i)	7.3	(1.7)	2.0	(0.0)	100.0	38.1	(0.0)	100.0	0.0
	split (c)	(ii)	3.7	(1.0)	3.0	(0.0)	100.0	33.3	(0.1)	100.0	0.0
		(iii)	1.0	(0.1)	2.0	(0.0)	100.0	16.1	(0.0)	100.0	0.0
Translog	pooled		4.4	(12.2)	2.0	(0.7)	98.1	108.9	(83.0)	81.7	7.8
		(i)	11.4	(32.4)	2.1	(0.9)	90.4	92.9	(76.4)	84.6	14.7
	split (a)	(ii)	7.9	(8.2)	1.9	(0.3)	100.0	26.4	(7.6)	100.0	0.0
		(iii)	2.0	(3.3)	1.9	(0.8)	100.0	42.0	(37.1)	68.5	10.2
		(i)	3.4	(9.8)	1.2	(0.8)	83.3	14.0	(38.3)	38.6	42.5
	split (c)	(ii)	2.7	(3.5)	1.6	(0.5)	100.0	69.1	(59.8)	80.8	0.0
		(iii)	2.3	(2.7)	2.3	(0.7)	100.0	75.1	(48.0)	80.0	11.7

(1) In column 3, 4 and 6, we report the mean of the corresponding variable over all observations of the (sub)sample. The sample standard deviation is given in parentheses. The lower and upper critical values for the null hypothesis of concavity are taken from Kodde and Palm [1986]. The critical values at the 5% threshold are given by $d_\ell = 2.706$ and $d_u = 10.371$.

(2) For the NQ, the concavity test should take the same value at any observation point. The finding of a positive standard deviation in column 6, may reveal the existence of several local minima to problem (8).

The second part of Table 3 (columns 6 to 9) reports some results on the statistical significance of concavity violations. As concavity is imposed locally at, say, $i = i^0$ and $t = t^0$, the result of our test depends on the arbitrary choice of the reference point (i^0, t^0). To avoid this inconvenience, we compute the test taking in turn each observation as reference point. The average level of $\widehat{d}$ over all reference points of the group and its sample standard deviation are gathered in column 6. The null hypothesis of concavity is rejected when the test statistic $\widehat{d}$ is found to be significantly different from zero. The percentage of cases for which $\widehat{d}$ was found to be significant (non significant) is reported in columns 7 and 8. As we use the upper and lower bounds proposed by Kodde and Palm [1986] for testing the inequality restrictions, these percentages do not sum up to 100, and it is not possible to reach a conclusion for every value of $\widehat{d}$.

The test points out that concavity violation is significant on average. For 18 out of 21 models the number of conclusive rejections exceeds that of conclusive failures to reject. In three cases (the BC on sub-sample (c) (i), the NQ on pooled data, and the TL on sub-sample (c) (i)), there are many test statistics falling in the inconclusive area (i.e. $d_\ell < \widehat{d} < d_u$), so that no conclusion can be reached.

What can be learned from this inference? First, no relationship between the frequency of concavity rejection and the number of degrees of freedom entailed in the model appears to exist: concavity is rejected (or not) independently of the sample split or functional forms considered. Second, whereas the tests of local concavity only provide weak evidence for a rejection, global concavity would be unambiguously rejected.

7. Restricted and unrestricted elasticities

In order to better understand the consequences of imposing concavity, we now compare restricted and unrestricted estimates of price, output and time elasticities. We also use the computed elasticities in order to study which model performs best in predicting the observed evolution of labor demand.

7.1 Own price elasticities

Tables 4 and 5 present own-price elasticities derived from models with and without concavity restriction, on the pooled sample and on sample split (a). Since the variations over time are not substantial, all elasticities are evaluated at 1985 values. To save space, we report the median value of each elasticity over the branches and its estimated standard error (s.e.) using the delta method. The concavity restriction is imposed for the year 1985 and for the branch producing the median level of output (No. 42).

For the pooled model, the concavity unrestricted results are not always plausible. With the BC and the TL, the own price elasticity of energy is significantly positive. With the NQ, the unrestricted own price elasticities have the expected sign but are always lower in absolute value than with the BC and TL. There is a great variability of the elasticities with respect to the functional form retained: ϵ_{hp_h} ranges between -1.39 for the TL and -0.37 for the NQ, and ϵ_{ep_e} ranges between -0.08 for the NQ and 0.88 for the TL. Note that inputs with a low cost share (as energy and high-skilled labor) have particularly variable own-price elasticities across the specifications. For material inputs, the own-price elasticity is rather stable.

When concavity is locally imposed on the estimates, all own-price elasticities become negative. For the observation at which concavity is imposed, the concavity test statistic $\widehat{d}$ is 79.7 for the BC, 9.5 for the NQ and 296.1 for the TL respectively. It can be seen that the restricted and unrestricted results are increasingly different with the importance of $\widehat{d}$. For the NQ specification, there are only small differences between restricted and unrestricted estimates, whereas for the TL model, the restricted ϵ_{hp_h} becomes implausible. This may explain why Diewert and Wales [1987] find relatively small differences between the unrestricted and restricted estimates, whereas Gagné and Ouellette [1998] show that the imposition of concavity may lead to important disparities between unrestricted and restricted estimates. Our main result is that concavity restrictions do not seem to be very useful with this type of data. In the single case where concavity cannot be statistically rejected (with the NQ), the restricted and unrestricted values are relatively similar. With the other two functional forms, concavity is rejected and the restricted own price elasticities are dramatically changed. With the BC, the imposition of concavity strongly affects the own-price elasticities for capital, high-skilled labor and energy. Coincidently these are the inputs with the smallest cost shares. With the TL this effect is even more pronounced.

The own-price elasticities obtained from the three sub-samples (split (a)) are reported in Table 5. For the BC, the value of the unrestricted own-price elasticity for energy ϵ_{ep_e} is now more plausible than the one obtained on the pooled sample. There is however some loss of precision in the BC estimates, since only 3 out of 6 own-price elasticities are significantly negative. Again, the own-price elasticities are smaller in absolute values for the NQ specification.

For the observation at which concavity is imposed, the concavity test statistic $\widehat{d}$ is 64.8 for the BC, $\widehat{d} = 1.9$ for the NQ and $\widehat{d} = 293.6$ for the TL. In this light it is not surprising that the disparity between the concavity unrestricted and restricted estimates are the most important for the TL specification. Contrary to the result obtained on the pooled sample, however, there is no huge difference between the restricted and the unrestricted elasticities. In fact, when the split model is restricted to fulfill concavity, only the estimates for one subsample are affected (the subsample containing the point (i^0, t^0)); the impact of concavity on the median elasticity is therefore limited. Note that the median of the concavity restricted elasticities ϵ_{mp_m} and ϵ_{ep_e} is not always negative in Table 5: as concavity is imposed at a given observation, the median elasticity may violate concavity.

The ranking of the own price elasticities of labor suggests that the demand for skilled labor is more elastic than for the demand for low-skilled labor. This stands in contradiction with findings in most previous studies (see Hamermesh 1993). Given the disparities that we have found, we must conclude that estimates of the own-price elasticities are highly

Table 4: Own-price elasticities, pooled data[1]

	Box-Cox			Normalized quadratic			Translog		
	unrestricted		restricted	unrestricted		restricted	unrestricted		restricted
	median	s.e.	median	median	s.e.	median	median	s.e.	median
ϵ_{kp_k}	−0.104	0.066	−0.204	−0.052	0.023	−0.063	−0.002	0.071	−0.392
ϵ_{hp_h}	−1.117	0.352	−1.606	−0.373	0.182	−0.618	−1.390	0.531	−3.277
ϵ_{sp_s}	−0.501	0.068	−0.522	−0.222	0.029	−0.242	−0.437	0.070	−0.565
ϵ_{up_u}	−0.458	0.091	−0.530	−0.241	0.112	−0.284	−0.148	0.136	−0.448
ϵ_{ep_e}	0.347	0.083	−0.074	−0.076	0.029	−0.090	0.878	0.130	−0.019
ϵ_{mp_m}	−0.021	0.019	−0.042	−0.013	0.013	−0.016	−0.061	0.024	−0.113

[1] Median value of the elasticities evaluated at the 1985 data and estimated standard error (s.e.).

Table 5: Own-price elasticities, sample split (a)[1]

	Box-Cox			Normalized quadratic			Translog		
	unrestricted		restricted	unrestricted		restricted	unrestricted		restricted
	median	s.e.	median	median	s.e.	median	median	s.e.	median
ϵ_{kp_k}	−0.027	0.030	−0.027	−0.035	0.007	−0.035	0.052	0.048	−0.152
ϵ_{hp_h}	−0.631	0.694	−0.615	−0.351	0.105	−0.353	−0.648	0.168	−1.321
ϵ_{sp_s}	−0.516	0.084	−0.445	−0.089	0.026	−0.092	−0.399	0.065	−0.569
ϵ_{up_u}	−0.387	0.099	−0.414	0.007	0.096	−0.057	−0.209	0.111	−0.752
ϵ_{ep_e}	0.003	0.076	−0.127	−0.101	0.042	−0.101	0.036	0.061	−0.278
ϵ_{mp_m}	0.010	0.068	0.010	0.015	0.004	0.015	−0.014	0.074	−0.014

[1] Median value of the elasticities over the three subsamples, evaluated at the 1985 data and estimated standard error (s.e.).

sensitive to the choice of functional form and sample split.[9]

7.2 Cross-price elasticities

To measure factor substitution possibilities, we compute cross-price elasticities for the unrestricted and concavity restricted models. The results for the pooled sample are given in Table D1, whereas these for sample split (a) are given in Table D2 in Appendix D. For the pooled sample BC unrestricted model, there are 15 out of 30 median cross-price elasticities which are significant at the 5 percent level. This number reduces to 9 for the NQ and to 12 for the TL cases. The number of significant elasticities obtained from the split sample are respectively 13, 14 and 13 for the BC, NQ and TL.

The cross-price elasticities computed on the basis of the NQ tend to be small in absolute value, pointing out a rigid production structure precluding frictionless substitution from one input against the other. Only 4 to 6 out of 30 elasticities are greater than 0.1 in absolute value for the NQ, depending on which sample and restrictions are chosen. For

[9] In an earlier version of the paper we found that the own-price elasticities (in absolute terms), estimated on the basis of the NQ specification, are decreasing with the level of skill. The difference in the results can be explained by a different estimation sample (only 27 industries have been retained) and an alternative definition of the user costs of capital.

the BC, there are 8 to 12 elasticities which are greater than 0.1 (respectively 12 to 14 for the TL). In comparison to our preferred specification (the BC), the TL overestimates and the NQ underestimates the extent of substitution and complementarity relationships.

Although there are some differences between the alternative models, the number of contradictions is not very important (we say that a contradiction occurs when an elasticity which is significantly different from zero in one model changes its sign or becomes insignificant in the other model). It can be observed that when concavity is not statistically rejected, the concavity adjusted elasticities do not differ much from the unrestricted ones (see the NQ case). The choice of the functional form and sample split has a greater impact on the estimates than the choice of whether to impose concavity or not.

We also observe some stable results for the elasticities of substitution. First, there is a dominant substitutability relationship between the three types of labor inputs: high-skilled and skilled labor can easily be substituted as well as skilled and unskilled labor. High-skilled labor cannot be substituted with any other input, and is complementary to unskilled labor. For all specifications considered, capital and energy are substitutes: a similar result is found in most previous studies for the US and Canada (see Thompson and Taylor, 1995). For some models, there is evidence for capital-skill complementarity, but this result is not robust with respect to the choice of functional form and does not hold in our preferred specification.[10] The results we have found are somewhat different to those found by Falk and Koebel [2000] and by Fitzenberger [1999] in the case where capital is quasi-fixed.

7.3 Output and time elasticities

Tables D3 and D4 (see Appendix D) present the output and time elasticities obtained from the pooled and split samples. These elasticities are, in most cases, significant at the five percent level. The results do not vary much across the specifications considered, and remain almost unaffected by the imposition of concavity. The main regularities are that: (i) there are increasing returns to scale ($\epsilon_{cy} \leq 1$); (ii) costs are reduced as time goes by ($\epsilon_{ct} \leq 0$); (iii) no input is regressive (or inferior), the elasticity of capital with respect to y is the lowest and the material-output elasticity is approximately equal to one; (iv) time is high-skill labor using, less-skilled labor saving ($\epsilon_{ht} \geq 0 \geq \epsilon_{st} \geq \epsilon_{ut}$), energy saving and material using. However, the interpretation of the time elasticities is delicate: they may pick up the influence of technical progress, but also the impact of any other omitted relevant variable which is correlated with time.

There are however some differences across the estimates. With the NQ, one would typically conclude that the output elasticity for different types of labor is increasingly positive with rising skill levels ($\epsilon_{hy} \geq \epsilon_{sy} \geq \epsilon_{uy}$). This result does not hold with the BC

[10] As in our model capital is assumed to be flexible, we adapt the definition given by Bergström and Panas [1992] and speak of capital-skill complementary when $\epsilon_{hp_k} \leq \epsilon_{sp_k} \leq \epsilon_{up_k}$. When $\epsilon_{up_k} \leq 0$, these inequalities mean that the degree of complementarity between labor and capital increases with skills. When $0 \leq \epsilon_{hp_k}$ it means that the degree of substitutability between labor and capital decreases with skills.

and the TL on the pooled sample. On sample split (a), there is some weak evidence for this hypothesis: for all functional forms retained $\epsilon_{hy} \geq \epsilon_{sy}$ and $\epsilon_{hy} \geq \epsilon_{uy}$. For all models considered, no contradiction can be found between the unrestricted and concavity restricted estimates.

7.4 Decomposition of factor demand growth

In order to better assess the performance of the different models, we now consider how well they can explain the observed shift away from unskilled labor and towards skilled labor that occurred over the period. For this purpose, one possibility would be to compare observed and predicted values of input demands for each specification. It is clear, from the statistical tests above, that the Box-Cox model on sample split (a) is the specification providing the best overall fit. As we are rather interested in the plausibility of the elasticities presented in Tables 4, 5 and D4 to D7, we follow an alternative approach in this sub-section, and study how well the evolution of input demands can be predicted using the alternative elasticities. For this purpose we decompose the predicted change in labor demand into three components reflecting the impact of factor substitution, growth and time respectively. These effects can be identified from the total differentiation of the labor demand equations:

$$\begin{aligned} \Delta g^*_{it} &\simeq \sum_{j=k,h,s,u,e,m} \frac{\partial g^*}{\partial p_{jit}} \Delta p_{jit} + \frac{\partial g^*}{\partial y_{it}} \Delta y_{it} + \frac{\partial g^*}{\partial t} \qquad (21) \\ \Leftrightarrow \quad \frac{\Delta g^*_{it}}{g^*_{it}} &\simeq \sum_{j=k,h,s,u,e,m} \varepsilon_{gp_j} \frac{\Delta p_{jit}}{p_{jit}} + \varepsilon_{gy} \frac{\Delta y_{it}}{y_{it}} + \varepsilon_{gt}, \end{aligned}$$

where $\Delta g^*_{it}/g^*_{it}$ denotes the predicted percentage change for the three types of labor ($g^*_{it} = h^*_{it}, s^*_{it}, u^*_{it}$). The observed values of the growth rates $\Delta g_{it}/g_{it}$, $\Delta p_{jit}/p_{jit}$ and $\Delta y_{it}/y_{it}$ can be calculated easily for each branch and time period. The different elasticities involved in (21) are computed for each branch and time period from the estimates. Then we compare the predicted and observed values ($\Delta g^*_{it}/g^*_{it}$ and $\Delta g_{it}/g_{it}$) for each branch and time period.

The first term on the right side of (21) measures the effect of own-price variation and input substitution, the second term reflects the impact of changes in the level of output and the last term denotes the impact of time. Note that the above decomposition is based on a *first order* approximation, and is only precise for small changes Δp_{jit} and Δy_{it}. Whereas a second order approximation would be more precise, the *separate identification* of the impact of price, output and time would then no longer be possible, as the second order terms involve interacting variables.

Columns three and five of Table 6 give the observed and predicted change for the three types of labor. In general, the predicted changes are relatively close to the observed ones. For instance, the (median) increase in the level of high-skilled labor is 3.2 percent which is close to the prediction of 2.9 percent with the BC. From the comparisons of predicted and observed values across functional forms, we can conclude that the Box-Cox seems to be the most reliable model: in 8 out of 12 cases, the evolution of h, s and u as explained by the BC is nearer to the observations than for the TL and NQ. The NQ appears to be the worst functional form, as it is never more precise than the BC or the TL.

Comparisons between concavity restricted and unrestricted specifications and between pooled and disaggregate models using the above criteria are inconclusive. In several cases the concavity restricted model does better than the unrestricted one, but as the differences are very small this does not warrant a firm conclusion. Concerning the estimates on sample split (a), they provide a better prediction for labor demands h and u, but lead to a worse prediction for skilled workers, independently of the functional form retained.

The last three columns of Table 6 show the decomposition (21). In general the impact of output and time is more important than price effects in explaining the shift towards skilled and away from unskilled labor. The impact of the evolution of prices is (almost) always negative, but also very small in absolute value (especially for the NQ). High-skilled labor is the labor input which is the most affected by the evolution of prices, due to the relatively high own-price elasticities. This suggests that wage pressure is an almost negligible factor for explaining the shift away from unskilled labor. For instance, only between 0 and 10 percent of the shift against unskilled labor can be explained by price effects. For h, the negative price effect is netted out by a positive impact of output growth, so that in the last instance the impact of time determines the overall evolution of high-skilled demand over the period. For skilled labor, output has the largest impact, and this widely offsets the negative effect of time. For unskilled labor, output is important too (at least for the BC and the TL), but it is largely outweighed by the impact of time.

Table 6: Determinants of labor demand by skill classes(1)

Functional form	input demand	actual change	modelling assumptions(2)	predicted change	% change attributable to: Price	Output	Time
Box-Cox	h	3.19	(a) / unr.	2.94	−0.76	0.93	2.76
			(a) / res.	3.10	−0.57	0.93	2.73
			(p) / unr.	2.19	−0.69	0.80	2.09
			(p) / res.	2.32	−0.66	0.72	2.26
	s	0.30	(a) / unr.	0.61	−0.15	1.15	−0.39
			(a) / res.	0.59	−0.09	1.09	−0.41
			(p) / unr.	0.27	−0.19	1.09	−0.63
			(p) / res.	0.43	−0.08	1.06	−0.55
	u	−3.19	(a) / unr.	−3.25	−0.20	1.32	−4.37
			(a) / res.	−3.26	−0.20	1.27	−4.33
			(p) / unr.	−3.43	0.03	1.24	−4.70
			(p) / res.	−3.39	0.08	1.11	−4.58
Normalized quadratic	h	3.19	(a) / unr.	2.46	−0.31	1.14	1.63
			(a) / res.	2.46	−0.31	1.14	1.63
			(p) / unr.	1.57	−0.15	0.89	0.83
			(p) / res.	1.80	−0.15	0.96	1.00
	s	0.30	(a) / unr.	0.97	−0.03	1.12	−0.12
			(a) / res.	0.98	−0.03	1.12	−0.12
			(p) / unr.	0.59	−0.01	0.92	−0.32
			(p) / res.	0.65	−0.01	0.93	−0.27
	u	−3.19	(a) / unr.	−2.94	0.03	0.05	−3.02
			(a) / res.	−2.94	0.02	0.05	−3.00
			(p) / unr.	−2.49	−0.02	0.05	−2.51
			(p) / res.	−2.44	−0.02	0.05	−2.47
Translog	h	3.19	(a) / unr.	2.79	−1.15	1.05	2.90
			(a) / res.	2.90	−1.06	1.05	2.92
			(p) / unr.	2.48	−1.38	1.27	2.59
			(p) / res.	2.44	−1.62	1.06	3.00
	s	0.30	(a) / unr.	0.73	−0.37	1.39	−0.29
			(a) / res.	0.93	−0.17	1.37	−0.27
			(p) / unr.	0.32	−0.31	1.23	−0.60
			(p) / res.	0.57	−0.21	1.23	−0.45
	u	−3.19	(a) / unr.	−2.97	−0.27	1.64	−4.34
			(a) / res.	−3.05	−0.27	1.62	−4.40
			(p) / unr.	−3.40	−0.05	1.73	−5.08
			(p) / res.	−3.42	−0.05	1.73	−5.09

(1) Column 3 shows the median growth rate over all industries and years. Columns 6 to 8 show the median value of the estimated impacts of price, output and time over all industries and years. The entries of column 5 are the sum of the corresponding entries of column 6 to 8.

(2) The letter (a) denotes sample split (a), (p) denotes the pooled sample, unr. stands for the concavity unrestricted specification and res. for the concavity restricted specification.

8. Conclusion

The purpose of this paper is to propose a method for imposing curvature conditions on a wide class of functional forms and for testing these restrictions. In the empirical application we estimate the parameters of a concavity constrained Box-Cox cost function. For our dataset, a parametric test for the null of concavity leads to a weak rejection of this assumption.

As concavity rejection may be related to a bad specification of functional form and to the heterogeneity of the observations, we also compare the performance of alternative model specifications and find that indeed the choice of functional form and sample split are important issues for obtaining plausible results. In particular the NQ functional form seems to underestimate the scope of substitution and complementary patterns. No relation could be found between the specification of the model and the frequency of concavity rejection. This may be related to the fact that the true aggregate relationships do not necessarily inherit all microeconomic properties (Koebel [2002]).

Concerning the determinants of labor demand, in general the impact of output and time is more important than price and substitution effects. We find that substitutability dominates between high skilled and skilled labor and between skilled and unskilled labor. Some complementarity is found between high-skilled and unskilled labor. The impact of prices and wages can, however, not explain much of the observed changes in the different types of labor inputs: while the evolution of skilled labor demand is mainly explained by output growth, the dominant factor 'explaining' the shift against unskilled labor and towards high-skilled labor is the residual time trend. This emphasizes the necessity to extend the usual theoretical framework in production analysis, in order to provide a better understanding of technological change and its determinants.

Appendix A: Proofs of the propositions

Proof of Proposition 1: For the sake of completeness, we adapt the results of Gouriéroux, Holly and Monfort [1982], Gouriéroux and Monfort [1989] and Wolak [1989] to prove the assertions of Proposition 1. For this aim, we also adopt their assumptions, which are not stated here for brevity.

Points (i) and (ii). It is clear that $\widehat{\alpha}$ and $\widetilde{\alpha}^0$ converges to the true parameter α_0 under H_0. It remains to show that $\widehat{\alpha}^0$ also converge to α_0. For this purpose, we characterize (in stages 1 to 3) the necessary conditions for an optimum in (8), (9) and (11), and show in stage 4, that the conditions corresponding to (8) are equivalent to those corresponding to (9) and (11).

Stage 1. Problem (8) can be reparameterized using the Cholesky decomposition in order to transform the inequality constraints $v'\overline{H}v \leq 0$, $\forall v$, into the *equality constraints* $g(\alpha) = \eta_H^0(u)$. The corresponding Lagrangean is

$$\mathcal{L} = (\widehat{\alpha} - \alpha)' \widehat{\Omega}_\alpha^{-1} (\widehat{\alpha} - \alpha) + \mu' \left(\eta_H^0(u) - g(\alpha)\right),$$

where $\mu \left(S_{\eta_H} \times 1\right)$ denotes the vector of Lagrange multipliers. The solution $\left(\widetilde{\alpha}^0, \widetilde{\mu}^0, \widetilde{u}^0\right)$ satisfies the first order conditions:

$$\frac{\partial \mathcal{L}}{\partial \alpha} = 0 \Leftrightarrow -2\widehat{\Omega}_\alpha^{-1}\left(\widehat{\alpha} - \widetilde{\alpha}^0\right) - \frac{\partial g'}{\partial \alpha}\left(\widetilde{\alpha}^0\right)\widetilde{\mu}^0 = 0 \quad \text{(A-1)}$$

$$\frac{\partial \mathcal{L}}{\partial u} = 0 \Leftrightarrow \frac{\partial \eta_H^{0\prime}}{\partial u}\left(\widetilde{u}^0\right)\widetilde{\mu}^0 = 0 \quad \text{(A-2)}$$

$$\frac{\partial \mathcal{L}}{\partial \mu} = 0 \Leftrightarrow \eta_H^0\left(\widetilde{u}^0\right) = g\left(\widetilde{\alpha}^0\right). \quad \text{(A-3)}$$

From $\partial\mathcal{L}/\partial\alpha = 0$ we obtain:

$$\widetilde{\mu}^0 = -2\left(\frac{\partial g}{\partial \alpha'}\left(\widetilde{\alpha}^0\right)\widehat{\Omega}_\alpha \frac{\partial g'}{\partial \alpha}\left(\widetilde{\alpha}^0\right)\right)^{-1} \frac{\partial g}{\partial \alpha'}\left(\widetilde{\alpha}^0\right)\left(\widehat{\alpha} - \widetilde{\alpha}^0\right) \quad \text{(A-4)}$$

and

$$\left(\widehat{\alpha} - \widetilde{\alpha}^0\right)' \widehat{\Omega}_\alpha^{-1}\left(\widehat{\alpha} - \widetilde{\alpha}^0\right) = \frac{\widetilde{\mu}^{0\prime}}{2}\frac{\partial g}{\partial \alpha'}\left(\widetilde{\alpha}^0\right)\widehat{\Omega}_\alpha \frac{\partial g'}{\partial \alpha}\left(\widetilde{\alpha}^0\right)\frac{\widetilde{\mu}^0}{2}. \quad \text{(A-5)}$$

Stage 2. The Lagrangean corresponding to (11) is given by

$$\mathcal{L} = (\widehat{\alpha} - \alpha)' \widehat{\Omega}_\alpha^{-1} (\widehat{\alpha} - \alpha) + \mu' \left(\widehat{\eta}_H^0 - g(\alpha)\right),$$

and the solution $\left(\widehat{\alpha}^0, \widehat{\mu}^0\right)$ satisfies:

$$\frac{\partial \mathcal{L}}{\partial \alpha} = 0 \Leftrightarrow -2\widehat{\Omega}_\alpha^{-1}\left(\widehat{\alpha} - \widehat{\alpha}^0\right) - \frac{\partial g'}{\partial \alpha}\left(\widehat{\alpha}^0\right)\widehat{\mu}^0 = 0 \quad \text{(A-6)}$$

$$\frac{\partial \mathcal{L}}{\partial \mu} = 0 \Leftrightarrow \widehat{\eta}_H^0 - g\left(\widehat{\alpha}^0\right) = 0. \quad \text{(A-7)}$$

Similarly to the former problem, we obtain

$$\widehat{\mu}^0 = -2\left(\frac{\partial g}{\partial \alpha'}\left(\widehat{\alpha}^0\right)\widehat{\Omega}_\alpha \frac{\partial g'}{\partial \alpha}\left(\widehat{\alpha}^0\right)\right)^{-1} \frac{\partial g}{\partial \alpha'}\left(\widehat{\alpha}^0\right)\left(\widehat{\alpha} - \widehat{\alpha}^0\right), \quad \text{(A-8)}$$

$$\left(\widehat{\alpha} - \widehat{\alpha}^0\right)' \widehat{\Omega}_\alpha^{-1}\left(\widehat{\alpha} - \widehat{\alpha}^0\right) = \frac{\widehat{\mu}^{0\prime}}{2}\frac{\partial g}{\partial \alpha'}\left(\widehat{\alpha}^0\right)\widehat{\Omega}_\alpha \frac{\partial g'}{\partial \alpha}\left(\widehat{\alpha}^0\right)\frac{\widehat{\mu}^0}{2}. \quad \text{(A-9)}$$

Stage 3. The first order conditions for a solution to (9) lead to:

$$\frac{\partial \eta_H^{0\prime}}{\partial u}\left(\widehat{u}^0\right)\widehat{\Omega}_H^{-1}\left(\widehat{\eta}_H-\eta_H^0\left(\widehat{u}^0\right)\right)=0, \tag{A-10}$$

Inserting (A-7) into (A-10) and after first order Taylor development around $\widehat{\alpha}$, this last condition becomes

$$\frac{\partial \eta_H^{0\prime}}{\partial u}\left(\widehat{u}^0\right)\widehat{\Omega}_H^{-1}\frac{\partial g}{\partial \alpha'}\left(\widehat{\alpha}\right)\left(\widehat{\alpha}-\widehat{\alpha}^0\right)\overset{a}{=}0,$$

which is asymptotically equivalent to (using (A-8) and (10))

$$\frac{\partial \eta_H^{0\prime}}{\partial u}\left(\widehat{u}^0\right)\widehat{\mu}^0\overset{a}{=}0. \tag{A-11}$$

Stage 4. In summary we have shown that the system (A-6)-(A-7), where $\widehat{\eta}_H^0\equiv\eta_H^0\left(\widehat{u}^0\right)$ is determined in (A-10), can be asymptotically equivalently written as

$$\begin{aligned}-2\widehat{\Omega}_\alpha^{-1}\left(\widehat{\alpha}-\widehat{\alpha}^0\right)-\tfrac{\partial g'}{\partial \alpha}\left(\widehat{\alpha}^0\right)\widehat{\mu}^0 &= 0\\ \eta_H^0\left(\widehat{u}^0\right)-g\left(\widehat{\alpha}^0\right) &= 0\\ \frac{\partial \eta_H^{0\prime}}{\partial u}\left(\widehat{u}^0\right)\widehat{\mu}^0 &\overset{a}{=} 0.\end{aligned}$$

This system comprises the same equations and unknown as the system (A-1)-(A-3), and as their solutions are unique they must be asymptotically identical.

Point (iii). As $\widehat{\alpha}$, $\widetilde{\alpha}^0$ and $\widehat{\alpha}^0$ converge to the true parameter α_0 under H_0 (Point i), we directly see from (A-4)-(A-5) and (A-8)-(A-9) of Point (i), that the minima of (8) and (11) are asymptotically equivalent. We now show that $\left(\widehat{\alpha}-\widehat{\alpha}^0\right)'\widehat{\Omega}_\alpha^{-1}\left(\widehat{\alpha}-\widehat{\alpha}^0\right)$ is asymptotically equivalent to $\left(\widehat{\eta}_H-\widehat{\eta}_H^0\right)'\widehat{\Omega}_H^{-1}\left(\widehat{\eta}_H-\widehat{\eta}_H^0\right)$ under H_0, with $\widehat{\eta}_H^0\equiv\eta_H^0\left(\widehat{u}^0\right)=g\left(\widehat{\alpha}^0\right)$. Using a first order Taylor development of $g\left(\widehat{\alpha}^0\right)$ around $\widehat{\alpha}$, we can write under H_0:

$$\begin{aligned}&\left[g\left(\widehat{\alpha}\right)-g\left(\widehat{\alpha}^0\right)\right]'\left[\frac{\partial g}{\partial \alpha'}\left(\widehat{\alpha}\right)\widehat{\Omega}_\alpha\frac{\partial g'}{\partial \alpha}\left(\widehat{\alpha}\right)\right]^{-1}\left[g\left(\widehat{\alpha}\right)-g\left(\widehat{\alpha}^0\right)\right]\\ \overset{a}{=}\;&\left(\widehat{\alpha}-\widehat{\alpha}^0\right)'\frac{\partial g'}{\partial \alpha}\left(\widehat{\alpha}^0\right)\left[\frac{\partial g}{\partial \alpha'}\left(\widehat{\alpha}\right)\widehat{\Omega}_\alpha\frac{\partial g'}{\partial \alpha}\left(\widehat{\alpha}\right)\right]^{-1}\frac{\partial g}{\partial \alpha'}\left(\widehat{\alpha}^0\right)\left(\widehat{\alpha}-\widehat{\alpha}^0\right)\\ =\;&\frac{\widehat{\mu}^{0\prime}}{2}\frac{\partial g}{\partial \alpha'}\left(\widehat{\alpha}^0\right)\widehat{\Omega}_\alpha\frac{\partial g'}{\partial \alpha}\left(\widehat{\alpha}^0\right)\left[\frac{\partial g}{\partial \alpha'}\left(\widehat{\alpha}\right)\widehat{\Omega}_\alpha\frac{\partial g'}{\partial \alpha}\left(\widehat{\alpha}\right)\right]^{-1}\frac{\partial g}{\partial \alpha'}\left(\widehat{\alpha}^0\right)\widehat{\Omega}_\alpha\frac{\partial g'}{\partial \alpha}\left(\widehat{\alpha}^0\right)\frac{\widehat{\mu}^0}{2}\\ \overset{a}{=}\;&\frac{\widehat{\mu}^{0\prime}}{2}\frac{\partial g}{\partial \alpha'}\left(\widehat{\alpha}^0\right)\widehat{\Omega}_\alpha\frac{\partial g'}{\partial \alpha}\left(\widehat{\alpha}^0\right)\frac{\widehat{\mu}^0}{2}\\ =\;&\left(\widehat{\alpha}-\widehat{\alpha}^0\right)'\widehat{\Omega}_\alpha^{-1}\left(\widehat{\alpha}-\widehat{\alpha}^0\right),\end{aligned}$$

where the fourth equality follows from points (i) and (ii) and the last equality follows from (A-9).

Point (iv). Using a first order Taylor expansion of (A-6) and (A-7) around $\widehat{\alpha}$, we can rewrite these conditions as

$$\begin{aligned}\widehat{\alpha}^0-\widehat{\alpha}\overset{a}{=}\widehat{\Omega}_\alpha\frac{\partial g}{\partial \alpha}\left(\widehat{\alpha}\right)'\frac{\widehat{\mu}^0}{2}\\ \widehat{\eta}_H^0-g\left(\widehat{\alpha}\right)-\frac{\partial g}{\partial \alpha'}\left(\widehat{\alpha}\right)\left(\widehat{\alpha}^0-\widehat{\alpha}\right)\overset{a}{=}0.\end{aligned}$$

Solving this system in $\left(\widehat{\alpha}^0, \widehat{\mu}^0\right)$ yields

$$\widehat{\mu}^0 \overset{a}{=} 2\left[\frac{\partial g}{\partial \alpha'}\left(\widehat{\alpha}\right)\widehat{\Omega}_\alpha \frac{\partial g'}{\partial \alpha}\left(\widehat{\alpha}\right)\right]^{-1}\left(\widehat{\eta}_H^0 - g\left(\widehat{\alpha}\right)\right),$$

and, finally:

$$\widehat{\alpha}^0 \overset{a}{=} \widehat{\alpha} + \widehat{\Omega}_\alpha \frac{\partial g'}{\partial \alpha}\left(\widehat{\alpha}\right)\left[\frac{\partial g}{\partial \alpha'}\left(\widehat{\alpha}\right)\widehat{\Omega}_\alpha \frac{\partial g'}{\partial \alpha}\left(\widehat{\alpha}\right)\right]^{-1}\left(\widehat{\eta}_H^0 - g\left(\widehat{\alpha}\right)\right).$$

Q.E.D.

Proof of Proposition 2: In order to ensure that the minimization in (13) occurs on the domain where $\lambda \leq 0$ we can simply reparameterize λ for imposing the nonpositivity of its components: for instance, define λ^0 as the vector with $-\nu_i^2$ as components. Then we can write

$$\begin{aligned} d_{KP} &= \min_{\lambda}\left\{\left(\widehat{\lambda}-\lambda\right)'\widehat{\Sigma}^-\left(\widehat{\lambda}-\lambda\right) : \lambda \leq 0\right\} \quad \text{(A-12)} \\ &= \min_{\nu}\left\{\left(\widehat{\lambda}-\lambda^0\left(\nu\right)\right)'\widehat{\Sigma}^-\left(\widehat{\lambda}-\lambda^0\left(\nu\right)\right)\right\} \\ &= \min_{\nu} \operatorname{vec}'\left(\widehat{\Lambda}-\Lambda^0\right)P'\widehat{\Sigma}^- P \operatorname{vec}\left(\widehat{\Lambda}-\Lambda^0\right), \end{aligned}$$

where P is defined as the selection matrix of size $S_p \times S_p^2$ such that $P\operatorname{vec}\left(\Lambda\right) = \lambda$. Let Q be the matrix of orthonormal eigenvectors of $\overline{H}$; then

$$Q'\overline{H}Q = \Lambda. \quad \text{(A-13)}$$

From this equation it follows that (see e.g. Magnus [1985] and Kodde and Palm [1987]):

$$\frac{\partial \lambda}{\partial \operatorname{vec}' \overline{H}} = P\left(Q' \otimes Q'\right).$$

Let $\widehat{\boldsymbol{\Omega}}_H$ be the (singular) variance matrix of $\operatorname{vec}\widehat{\overline{H}}$; the variance matrix considered by Kodde and Palm can then be written as

$$\begin{aligned} \widehat{\Sigma}^- &= \left[\frac{\partial \lambda}{\partial \operatorname{vec}'\left(\overline{H}\right)}\left(\widehat{\overline{H}}\right)\widehat{\boldsymbol{\Omega}}_H \frac{\partial \lambda'}{\partial \operatorname{vec}\left(\overline{H}\right)}\left(\widehat{\overline{H}}\right)\right]^- = \left[P\left(\widehat{Q}' \otimes \widehat{Q}'\right)\widehat{\boldsymbol{\Omega}}_H\left(\widehat{Q}' \otimes \widehat{Q}'\right)'P'\right]^- \\ &= P\left(\widehat{Q}' \otimes \widehat{Q}'\right)\widehat{\boldsymbol{\Omega}}_H^-\left(\widehat{Q}' \otimes \widehat{Q}'\right)'P', \quad \text{(A-14)} \end{aligned}$$

using the orthogonality of $\left(\widehat{Q}' \otimes \widehat{Q}'\right)$ and $\left(\widehat{Q}' \otimes \widehat{Q}'\right)'$ and the fact that PP' is an identity matrix.[11] Note that $\widehat{\Omega}_H$, the covariance matrix of $\widehat{\eta}_H \equiv \operatorname{vecli}\widehat{\overline{H}}$, is a submatrix of $\widehat{\boldsymbol{\Omega}}_H$. Let $\overline{H}^0 = \widehat{Q}\Lambda^0\widehat{Q}'$, using (A-12) and (A-14), we can rewrite

$$\begin{aligned} d_{KP} &= \min_{\nu} \operatorname{vec}'\left(\widehat{Q}'\widehat{\overline{H}}\widehat{Q}-\Lambda^0\right)P'\widehat{\Sigma}^- P\operatorname{vec}\left(\widehat{Q}'\widehat{\overline{H}}\widehat{Q}-\Lambda^0\right) \\ &= \min_{\nu} \operatorname{vec}'\left(\widehat{Q}'\left(\widehat{\overline{H}}-\overline{H}^0\right)\widehat{Q}\right)P'\widehat{\Sigma}^- P\operatorname{vec}\left(\widehat{Q}'\left(\widehat{\overline{H}}-\overline{H}^0\right)\widehat{Q}\right) \\ &= \min_{\nu} \operatorname{vec}'\left(\widehat{\overline{H}}-\overline{H}^0\right)\left(\widehat{Q} \otimes \widehat{Q}\right)P'\widehat{\Sigma}^- P\left(\widehat{Q}' \otimes \widehat{Q}'\right)\operatorname{vec}\left(\widehat{\overline{H}}-\overline{H}^0\right) \\ &= \min_{\nu} \operatorname{vec}'\left(\widehat{\overline{H}}-\overline{H}^0\right)\widehat{\boldsymbol{\Omega}}_H^-\operatorname{vec}\left(\widehat{\overline{H}}-\overline{H}^0\right) \\ &= \min_{\nu} \operatorname{vecli}'\left(\widehat{\overline{H}}-\overline{H}^0\right)\widehat{\Omega}_H^{-1}\operatorname{vecli}\left(\widehat{\overline{H}}-\overline{H}^0\right), \end{aligned}$$

[11] It is then easy to check that, denoting $A = P\left(\widehat{Q}' \otimes \widehat{Q}'\right)\widehat{\boldsymbol{\Omega}}_H\left(\widehat{Q}' \otimes \widehat{Q}'\right)'P'$, one has indeed $AA^-A = A$.

where the second equality follows from the properties of the vec operator and Kronecker product,[12] and the fourth from Dhrymes' [1994] Lemma A1.[13] As the matrix $\overline{H}^0 = \widehat{Q}\Lambda^0\widehat{Q}'$ is negative semi-definite, it can be written as $-U'U$. However, the matrix $\widehat{Q}\Lambda^0\widehat{Q}'$ comprises only the $S_p - 1$ free parameters of ν, whereas $-U'U$ is made of the $S_p(S_p-1)/2$ free parameters u. Hence in small samples

$$\begin{aligned} d_{KP} &= \min_{\nu} \operatorname{vecli}'\left(\widehat{\overline{H}} - \overline{H}^0\right)\widehat{\Omega}_H^{-1}\operatorname{vecli}\left(\widehat{\overline{H}} - \overline{H}^0\right) \\ &\geq \min_{u} \operatorname{vecli}'\left(\widehat{\overline{H}} + U'U\right)\widehat{\Omega}_H^{-1}\operatorname{vecli}\left(\widehat{\overline{H}} + U'U\right) \\ &= \min_{u} \left[\widehat{\eta}_H - \eta_H^0(u)\right]'\widehat{\Omega}_H^{-1}\left[\widehat{\eta}_H - \eta_H^0(u)\right] = d, \end{aligned}$$

with $\widehat{\eta}_H \equiv \operatorname{vecli}\widehat{\overline{H}}$ and $\eta_H^0(u) \equiv \operatorname{vecli}(-U'U)$. The reason for the inequality is that in the minimization over u, which entails more parameters than ν, can only reach a smaller value than in the case when the minimization is over ν. By minimizing over ν, one does not change the eigenvectors $\widehat{Q}$, whereas the minimization over u affects simultaneously eigenvalues and eigenvectors. Asymptotically $d_{KP} \overset{a}{=} d$, because under H_0, a first order Taylor development of $\widehat{\lambda}^0$ around $\widehat{\lambda}$ allows to write

$$\begin{aligned} &\left(\widehat{\lambda} - \widehat{\lambda}^0\right)'\widehat{\Sigma}^-\left(\widehat{\lambda} - \widehat{\lambda}^0\right) \\ \overset{a}{=}\ & \operatorname{vec}'\left(\widehat{\overline{H}} - \overline{H}^0\right)\frac{\partial\lambda}{\partial\operatorname{vec}\left(\overline{H}\right)}\left(\widehat{\overline{H}}\right)\widehat{\Sigma}^-\frac{\partial\lambda}{\partial\operatorname{vec}'\left(\overline{H}\right)}\left(\widehat{\overline{H}}\right)\operatorname{vec}\left(\widehat{\overline{H}} - \overline{H}^0\right) \\ =\ & \left(\widehat{\eta}_H - \widehat{\eta}_H^0\right)'\widehat{\Omega}_H^{-1}\left(\widehat{\eta}_H - \widehat{\eta}_H^0\right). \end{aligned}$$

Q.E.D.

Proof of Proposition 3: (i) An eigenvalue λ_1 of $\overline{H}$ is a solution of $f\left(\lambda, \overline{H}\right) \equiv \left|\lambda I_{S_p} - \overline{H}\right| = 0$.[14] This eigenvalue λ_1 can be expressed as a function of the parameters of $\overline{H}$ when the conditions of the implicit function theorem are fulfilled, that is, when $\partial f\left(\lambda, \overline{H}\right)/\partial\lambda \neq 0$ at $\lambda = \lambda_1$. Since

$$\left|\lambda I_{S_p} - \overline{H}\right| = \left|\lambda I_{S_p} - \Lambda\right| = \prod_{i=1}^{S_p}(\lambda - \lambda_i),$$

it follows that

$$\frac{\partial f\left(\lambda, \overline{H}\right)}{\partial\lambda} = \sum_{j=1}^{S_p}\prod_{i\neq j}^{S_p}(\lambda - \lambda_i).$$

Hence

$$\left.\frac{\partial f\left(\lambda, \overline{H}\right)}{\partial\lambda}\right|_{\lambda=\lambda_1} = \prod_{i\neq 1}^{S_p}(\lambda_1 - \lambda_i).$$

Thus, $\partial f\left(\lambda, \overline{H}\right)/\partial\lambda$ is different from zero if and only if λ_1 is simple. A related result is

[12] For three matrices of adequate dimensions, $\operatorname{vec}(ABC) = (C' \otimes A)\operatorname{vec}(B)$.

[13] Dhrymes' [1994] Lemma A1: Let $x \sim N(\mu, \Sigma)$ be a m -element (column) vector of random variables and suppose rank$(\Sigma) = r \leq m$. Let $x_{(1)}$ be an *arbitrary* subset of r elements of x, such that its covariance $\Sigma_{(11)}$ is *nonsingular*. Then, in the obvious notation, where Σ_g is the (unique) *generalized inverse*,

$$(x-\mu)'\Sigma_g(x-\mu) = \left(x_{(1)} - \mu_{(1)}\right)'\Sigma_{(11)}^{-1}\left(x_{(1)} - \mu_{(1)}\right)$$

Remark that the result is valid for any vector x (the normality assumption is not necessary).

[14] Notation λ is used here for a scalar, and no longer for the vector used in the proof of Proposition 2.

obtained by Magnus [1985, Theorem 1].
(ii) We adapt a proof given by Lau [1978, Lemmata 3.6 and 3.7] to our slightly different problem. Let $\mathcal{H}_{S_p-1}$ be the set of all real symmetric matrices with rank $S_p - 1$ and let $\mathcal{H}^0_{S_p-1}$ be the subset of $\mathcal{H}_{S_p-1}$ of all matrices with multiple eigenvalues. For given eigenvalues $\lambda_2, \ldots, \lambda_{S_p}$, the set of λ_1 such that $\prod_{i\neq 1}^{S_p}(\lambda_1 - \lambda_i) = 0$ is a set of measure zero. As the union (over $j = 1, \ldots, S_p$) of a countable number of null sets is again a null set, the subset $\mathcal{H}^0_{S_p-1}$ of matrices satisfying $\prod_{i\neq j}^{S_p}(\lambda_j - \lambda_i) = 0$, $j = 1, \ldots, S_p$, is of measure zero. *Q.E.D.*

Appendix B: Some functional forms nested within the BC

In this appendix we show how the normalized quadratic, a version of the generalized Leontief, and the translog, are obtained as special cases of the generalized Box-Cox specification. The derivation of further interesting functional forms can also be obtained along these lines.

- For $\gamma_1 = \gamma_2 = 1$, the normalized quadratic cost function is obtained as a special case of the Box-Cox specification. Indeed, we then have

$$\begin{aligned} Z_{it} &= z_{it} - \iota_{S_z}, \\ P_{it} &= \frac{p_{it}}{\theta_i' p_{it}} - \iota_{S_p}, \end{aligned}$$

where ι_{S_p} is a S_p-vector of ones. The cost function (14) then becomes

$$\begin{aligned} c^* &= \theta_i' p_{it} C^*_{NQ}(p_{it}, z_{it}, \alpha) + (\theta_i' p_{it}) \\ &= \theta_i' p_{it}\left(1 + \beta_{0i} + P_{it}' B_{pi} + Z_{it}' B_z + \tfrac{1}{2} P_{it}' B_{pp} P_{it} + P_{it}' B_{pz} Z_{it} + \tfrac{1}{2} Z_{it}' B_{zz} Z_{it}\right) \\ &= p_{it}' B_{pi} + (\theta_i' p_{it})\, z_{it}' B_z + \tfrac{1}{2}\frac{p_{it}' B_{pp} p_{it}}{\theta_i' p_{it}} + p_{it}' B_{pz} z_{it} + (\theta_i' p_{it})\, \tfrac{1}{2} z_{it}' B_{zz} z_{it} \\ &\quad + \theta_i' p_{it}\left(1 + \beta_{0i} - \iota_{S_p}' B_{pi} - \iota_{S_z}' B_z + \tfrac{1}{2}\iota_{S_p}' B_{pp}\iota_{S_p} + \iota_{S_p}' B_{pz}\iota_{S_z} + \tfrac{1}{2}\iota_{S_z}' B_{zz}\iota_{S_z}\right) \\ &\quad + \theta_i' p_{it}\left(-\iota_{S_p}' B_{pp}\frac{p_{it}}{\theta_i' p_{it}} - \iota_{S_p}' B_{pz} z_{it} - \frac{p_{it}'}{\theta_i' p_{it}} B_{pz}\iota_{S_z} - \iota_{S_z}' B_{zz} z_{it}\right) \end{aligned}$$

Considering the restrictions (16), we obtain

$$\begin{aligned} c^* &= p_{it}' B_{pi} + \theta_i' p_{it} z_{it}' B_{zi} + \tfrac{1}{2}\frac{p_{it}' B_{pp} p_{it}}{\theta_i' p_{it}} + p_{it}' B_{pz} z_{it} + \tfrac{1}{2}\theta_i' p_{it} z_{it}' B_{zz} z_{it} \\ &\quad + \theta_i' p_{it}\left(\beta_{0i} - \iota_{S_z}' B_z + \tfrac{1}{2}\iota_{S_z}' B_{zz}\iota_{S_z}\right) + \theta_i' p_{it}\left(-\frac{p_{it}'}{\theta_i' p_{it}} B_{pz}\iota_{S_z} - \iota_{S_z}' B_{zz} z_{it}\right) \\ &= p_{it}' \mathbf{B}_{pi} + \theta_i' p_{it} z_{it}' \mathbf{B}_z + \tfrac{1}{2}\frac{p_{it}' B_{pp} p_{it}}{\theta_i' p_{it}} + p_{it}' B_{pz} z_{it} + \tfrac{1}{2} z_{it}' B_{zz} z_{it}, \end{aligned}$$

which is the expression of the normalized quadratic cost function, with

$$\begin{aligned} \mathbf{B}_{pi} &= B_{pi} - B_{pz}\iota_{S_z} + \theta_i\left(\beta_{0i} - \iota_{S_z}' B_z + \tfrac{1}{2}\iota_{S_z}' B_{zz}\iota_{S_z}\right) \\ \mathbf{B}_z &= B_z - B_{zz}\iota_{S_z}. \end{aligned}$$

- In the case where $\gamma_1 = 1/2$ and $\gamma_2 = 1$, we have

$$\begin{aligned} Z_{it} &= 2\left(z_{it}^{1/2} - \iota_{S_z}\right), \\ P_{it} &= 2\left[\left(\frac{p_{it}}{\theta_i' p_{it}}\right)^{1/2} - \iota_{S_p}\right], \end{aligned}$$

where by convention $z_{it}^{1/2} = \left(z_1^{1/2}, \ldots, z_{S_z}^{1/2}\right)_{it}'$ and $p_{it}^{1/2} = \left(p_1^{1/2}, \ldots, p_{S_p}^{1/2}\right)_{it}'$. The cost function (14) then becomes

$$\begin{aligned} c^* &= \theta_i' p_{it} C_{GL}^* \left(p_{it}, z_{it}, \alpha\right) + \theta_i' p_{it} \\ &= \theta_i' p_{it} \left(1 + \beta_{0i} + P_{it}' B_{pi} + Z_{it}' B_z + \tfrac{1}{2} P_{it}' B_{pp} P_{it} + P_{it}' B_{pz} Z_{it} + \tfrac{1}{2} Z_{it}' B_{zz} Z_{it}\right) \\ &= 2\left(\theta_i' p_{it}\right)^{1/2} p_{it}^{1/2\prime} B_{pi} + 2\theta_i' p_{it} z_{it}^{1/2\prime} B_{zi} + 2 p_{it}^{1/2\prime} B_{pp} p_{it}^{1/2} \\ &\quad + 4\left(\theta_i' p_{it}\right)^{1/2} p_{it}^{1/2\prime} B_{pz} z_{it}^{1/2} + 2\theta_i' p_{it} z_{it}^{1/2\prime} B_{zz} z_{it}^{1/2} \\ &\quad + 2\theta_i' p_{it} \left(1 + \beta_{0i} - \iota_{S_p}' B_{pi} - \iota_{S_z}' B_z + \iota_{S_p}' B_{pp} \iota_{S_p} + 2\iota_{S_p}' B_{pz} \iota_{S_z} + \iota_{S_z}' B_{zz} \iota_{S_z}\right) \\ &\quad + 4\theta_i' p_{it} \left[-\iota_{S_p}' B_{pp} \left(\frac{p_{it}}{\theta_i' p_{it}}\right)^{1/2} - \iota_{S_p}' B_{pz} z_{it}^{1/2} - \left(\frac{p_{it}'}{\theta_i' p_{it}}\right)^{1/2} B_{pz} \iota_{S_z} - \iota_{S_z}' B_{zz} z_{it}^{1/2}\right] \\ &= 2\left(\theta_i' p_{it}\right)^{1/2} p_{it}^{1/2\prime} B_{pi} + 2\theta_i' p_{it} z_{it}^{1/2\prime} B_z + 2 p_{it}^{1/2\prime} B_{pp} p_{it}^{1/2} \\ &\quad + 4\left(\theta_i' p_{it}\right)^{1/2} p_{it}^{1/2\prime} B_{pz} z_{it}^{1/2} + 2\theta_i' p_{it} z_{it}^{1/2\prime} B_{zz} z_{it}^{1/2} \\ &\quad + 2\theta_i' p_{it} \left(\beta_{0i} - \iota_{S_z}' B_z + \iota_{S_z}' B_{zz} \iota_{S_z}\right) + 4\theta_i' p_{it} \left[-\left(\frac{p_{it}'}{\theta_i' p_{it}}\right)^{1/2} B_{pz} \iota_{S_z} - \iota_{S_z}' B_{zz} z_{it}^{1/2}\right]. \end{aligned}$$

After reparameterization, we obtain

$$\begin{aligned} c^* &= \left(\theta_i' p_{it}\right)^{1/2} p_{it}^{1/2\prime} \mathbf{B}_{pi} + \theta_i' p_{it} z_{it}^{1/2\prime} \mathbf{B}_z + p_{it}^{1/2\prime} \mathbf{B}_{ppi} p_{it}^{1/2} \\ &\quad + \left(\theta_i' p_{it}\right)^{1/2} p_{it}^{1/2\prime} \mathbf{B}_{pz} z_{it}^{1/2} + \theta_i' p_{it} z_{it}^{1/2\prime} \mathbf{B}_{zz} z_{it}^{1/2}, \end{aligned}$$

which is a version of the generalized Leontief cost function, with

$$\begin{aligned} \mathbf{B}_{pi} &= 2B_{pi} - 4B_{pz}\iota_{S_z}, \\ \mathbf{B}_z &= 2B_z - 4B_{zz}\iota_{S_z}, \\ \mathbf{B}_{ppi} &= 2B_{pp} + \operatorname{diag}\left[2\theta_i\left(\beta_{0i} - \iota_{S_z}' B_z + \iota_{S_z}' B_{zz}\iota_{S_z}\right)\right], \\ \mathbf{B}_{pz} &= 4B_{pz}, \\ \mathbf{B}_{zz} &= 2B_{zz}, \end{aligned}$$

where $\operatorname{diag}(v)$ is a diagonal matrix with the vector v on the main diagonal.

- When $\gamma_1 \rightarrow 0$ and $\gamma_2 \rightarrow 0$, the translog is obtained as a limiting case. Indeed,

$$\begin{aligned} Z_{it} &= \lim_{\gamma_1 \rightarrow 0} \frac{z_{it}^{\gamma_1} - 1}{\gamma_1} &&= \ln z_{it}, \\ C_{TL}^* &= \lim_{\gamma_2 \rightarrow 0} \frac{\left(c^*/\theta_i' p_{it}\right)^{\gamma_2} - 1}{\gamma_2} &&= \ln c^* - \ln\left(\theta_i' p_{it}\right), \\ P_{it} &= \lim_{\gamma_1 \rightarrow 0} \frac{\left(p_{it}/\theta_i' p_{it}\right)^{\gamma_1} - \iota_{S_p}}{\gamma_1} &&= \ln p_{it} - \iota_{S_p} \ln\left(\theta_i' p_{it}\right), \end{aligned}$$

where by convention $\ln z_{it} = \left(\ln z_1, \ldots, \ln z_{S_z}\right)_{it}'$ and $\ln p_{it} = \left(\ln p_1, \ldots, \ln p_{S_p}\right)_{it}'$. The cost function then becomes

$$
\begin{aligned}
\ln c^* &= C^*_{TL}\left(p_{it}, z_{it}, \alpha_i\right) + \ln\left(\theta_i' p_{it}\right) \\
&= \left(\beta_{0i} + P_{it}' B_{pi} + Z_{it}' B_{zi} + \tfrac{1}{2} P_{it}' B_{pp} P_{it} + P_{it}' B_{pz} Z_{it} + \tfrac{1}{2} Z_{it}' B_{zz} Z_{it}\right) + \ln\left(\theta_i' p_{it}\right) \\
&= \beta_{0i} + \left(\ln p_{it}\right)' B_{pi} + \left(\ln z_{it}\right)' B_{zi} + \tfrac{1}{2}\left(\ln p_{it}\right)' B_{pp}\left(\ln p_{it}\right) \\
&\quad + \left(\ln p_{it}\right)' B_{pz} \ln z_{it} + \tfrac{1}{2}\left(\ln z_{it}\right)' B_{zz} \ln z_{it} \\
&\quad + \ln\left(\theta_i' p_{it}\right)\left(1 - \iota_{S_p}' B_{pi} + \tfrac{1}{2} \ln\left(\theta_i' p_{it}\right) \iota_{S_p}' B_{pp} \iota_{S_p} - \iota_{S_p}' B_{pp}\left(\ln p_{it}\right) - \iota_{S_p}' B_{pz} \ln z_{it}\right)
\end{aligned} \tag{B-1}
$$

After restrictions (16) have been imposed, the last line of (B-1) vanishes and the usual translog specification is obtained.

Appendix C: Description of the branches

Denomination of the branches

No.	Branch	No.	Branch
14	Chemical products	30	Electrical machinery, equipment and appliances
15	Refined petroleum products	31	Precision instruments and optical equipment
16	Plastic products	32	Tools and finished metal products
17	Rubber products	33	Musical instruments, games and toys, jewelry, etc.
18	Quarrying, building materials, etc.	34	Wood working
19	Ceramic products	35	Wood products
20	Glass products	36	Pulp, paper and paperboard
21	Iron and steel	37	Paper and paperboard products
22	Non-ferrous metals, etc.	38	Paper processing
23	Foundry products	39	Printing and reproduction
24	Drawing plants products, cold rolling mills, etc.	40	Leather and leather products, footwear
25	Structural metal products, rolling stock	41	Textiles
26	Machinery and equipment	42	Wearing apparel
27	Office machinery and computers	43	Food products (excl. beverages)
28	Road vehicles	44	Beverages
29	Ships and boats	45	Tobacco products
30	Aircraft and spacecraft		

Description of the different sample splits considered

Split (a)	consumer goods	investment goods	intermediate goods
Industry No.	16, 19, 20, 34, 36, 38, 39, 40, 41, 42, 43, 44, 45.	25, 26, 27, 28, 29, 30, 31, 32, 33.	14, 17, 18, 21, 22, 23, 24, 35, 37.
Split (b)	small branches	medium branches	large branches
Industry No.	17, 19, 20, 23, 29, 30, 34, 35, 37, 38, 40.	18, 22, 25, 27, 32, 36, 39, 42, 44, 45.	14, 16, 21, 24, 26, 28, 31, 33, 41, 43.
Split (c)	low skill intensive	skill intensive	high skill intensive
Industry No.	24, 34, 35, 36, 38, 39, 40, 41, 42, 43, 44.	16, 18, 20, 21, 22, 23, 28, 33, 37, 45.	14, 17, 19, 25, 26, 27, 29, 30, 31, 32.
Split (d)	not labor intensive	labor intensive	highly labor intensive
Industry No.	14, 21, 22, 27, 35, 37, 38, 40, 43, 44, 45.	16, 17, 18, 20, 24, 28, 29, 34, 41, 42.	19, 23, 25, 26, 30, 31, 32, 33, 36, 39.
Split (e)	not capital intensive	capital intensive	highly capital intensive
Industry No.	16, 22, 25, 26, 31, 32, 33, 36, 40, 42, 43.	14, 17, 23, 24, 28, 30, 34, 38, 41, 45.	18, 19, 20, 21, 27, 29, 35, 37, 39, 44.

Appendix D: Further estimation results

Table D1: Cross price elasticities, pooled sample

	Box-Cox			Normalized quadratic			Translog		
	unrestricted		restricted	unrestricted		restricted	unrestricted		restricted
	median	s.e.	median	median	s.e.	median	median	s.e.	median
ϵ_{kp_h}	−0.116	0.023	−0.094	−0.070	0.014	−0.058	−0.132	0.028	−0.120
ϵ_{kp_s}	−0.091	0.039	−0.022	−0.005	0.009	0.016	−0.151	0.050	0.104
ϵ_{kp_u}	0.073	0.038	0.089	0.058	0.038	0.053	0.104	0.031	0.119
ϵ_{kp_e}	0.068	0.025	0.026	0.040	0.012	0.034	0.053	0.030	0.009
ϵ_{kp_m}	0.154	0.036	0.200	0.038	0.058	0.014	0.107	0.067	0.243
ϵ_{hp_k}	−0.770	0.157	−0.608	−0.399	0.087	−0.334	−0.914	0.179	−0.841
ϵ_{hp_s}	2.439	0.492	2.095	1.191	0.219	1.134	3.419	0.544	3.457
ϵ_{hp_u}	−0.264	0.372	0.296	−0.233	0.273	−0.013	−0.892	0.539	1.069
ϵ_{hp_e}	−0.054	0.112	−0.048	0.014	0.066	0.007	−0.037	0.116	−0.037
ϵ_{hp_m}	−0.143	0.287	−0.115	−0.130	0.117	−0.129	−0.138	0.309	−0.255
ϵ_{sp_k}	−0.045	0.020	−0.010	−0.002	0.008	0.010	−0.057	0.021	0.046
ϵ_{sp_h}	0.166	0.027	0.156	0.066	0.012	0.062	0.250	0.035	0.247
ϵ_{sp_u}	0.262	0.056	0.263	0.144	0.034	0.144	0.054	0.032	0.065
ϵ_{sp_e}	0.000	0.015	0.028	0.000	0.013	0.005	−0.006	0.035	0.017
ϵ_{sp_m}	0.099	0.052	0.091	0.031	0.016	0.043	0.152	0.038	0.127
ϵ_{up_k}	0.066	0.036	0.079	0.046	0.029	0.045	0.098	0.022	0.117
ϵ_{up_h}	−0.038	0.054	0.038	−0.029	0.027	−0.003	−0.118	0.071	0.146
ϵ_{up_s}	0.486	0.109	0.454	0.241	0.062	0.242	0.102	0.084	0.123
ϵ_{up_e}	0.023	0.019	0.024	−0.027	0.020	−0.017	0.020	0.021	0.008
ϵ_{up_m}	−0.081	0.040	−0.065	−0.007	0.026	−0.001	−0.040	0.088	−0.008
ϵ_{ep_k}	0.150	0.058	0.056	0.081	0.026	0.066	0.125	0.076	0.029
ϵ_{ep_h}	−0.033	0.035	−0.030	0.003	0.134	0.001	−0.023	0.036	−0.024
ϵ_{ep_s}	−0.002	0.076	0.120	0.000	0.047	0.028	−0.052	0.069	0.119
ϵ_{ep_u}	0.065	0.047	0.066	−0.059	0.046	−0.043	0.059	0.127	0.036
ϵ_{ep_m}	−0.541	0.134	−0.161	0.055	0.046	0.036	−0.992	0.182	−0.085
ϵ_{mp_k}	0.022	0.012	0.029	0.004	0.003	0.002	0.011	0.008	0.029
ϵ_{mp_h}	−0.002	0.004	−0.002	−0.003	0.004	−0.003	−0.003	0.006	−0.005
ϵ_{mp_s}	0.024	0.012	0.022	0.012	0.007	0.016	0.045	0.013	0.036
ϵ_{mp_u}	−0.011	0.010	−0.010	−0.001	0.014	0.000	−0.004	0.010	−0.001
ϵ_{mp_e}	−0.024	0.007	−0.006	0.004	0.003	0.002	−0.042	0.008	−0.004

(1) Median value of the elasticities evaluated at the 1985 data and estimated standard error (s.e.).

Table D2: Cross price elasticities, sample split (a)

	Box-Cox			Normalized quadratic			Translog		
	unrestricted		restricted	unrestricted		restricted	unrestricted		restricted
	median	s.e.	median	median	s.e.	median	median	s.e.	median
ϵ_{kp_h}	−0.001	0.008	0.007	−0.011	0.007	−0.011	−0.015	0.008	−0.008
ϵ_{kp_s}	−0.108	0.038	−0.052	−0.074	0.052	−0.074	−0.147	0.170	−0.078
ϵ_{kp_u}	0.004	0.046	0.032	0.037	0.015	0.044	−0.001	0.063	0.001
ϵ_{kp_e}	0.111	0.018	0.025	0.041	0.007	0.038	0.097	0.013	0.057
ϵ_{kp_m}	0.005	0.053	−0.021	0.034	0.013	0.034	0.001	0.218	0.068
ϵ_{hp_k}	−0.006	0.098	0.024	−0.020	0.006	−0.020	−0.099	0.162	−0.042
ϵ_{hp_s}	1.209	0.214	1.209	0.128	0.039	0.128	1.258	0.216	1.258
ϵ_{hp_u}	−0.484	0.339	−0.484	−0.009	0.076	0.027	−0.713	0.926	−0.471
ϵ_{hp_e}	−0.097	0.103	−0.097	0.017	0.073	0.015	−0.173	0.077	−0.031
ϵ_{hp_m}	0.112	0.232	0.112	0.280	0.138	0.280	0.130	0.127	0.130
ϵ_{sp_k}	−0.055	0.045	−0.027	−0.043	0.031	−0.043	−0.055	0.068	−0.035
ϵ_{sp_h}	0.124	0.027	0.124	0.021	0.005	0.021	0.152	0.035	0.152
ϵ_{sp_u}	0.159	0.065	0.153	−0.007	0.028	0.003	0.128	0.071	0.131
ϵ_{sp_e}	0.070	0.020	0.070	0.043	0.012	0.042	0.041	0.016	0.053
ϵ_{sp_m}	−0.003	0.041	−0.003	−0.012	0.008	−0.012	0.115	0.048	0.115
ϵ_{up_k}	0.003	0.045	0.024	0.034	0.014	0.034	−0.001	0.040	0.001
ϵ_{up_h}	−0.070	0.045	−0.070	−0.001	0.005	0.002	−0.077	0.122	−0.057
ϵ_{up_s}	0.269	0.106	0.251	−0.008	0.028	0.005	0.184	0.131	0.184
ϵ_{up_e}	−0.011	0.023	−0.008	0.002	0.008	0.002	0.028	0.016	0.028
ϵ_{up_m}	0.162	0.143	0.150	−0.082	0.027	−0.082	0.166	0.044	0.166
ϵ_{ep_k}	0.167	0.060	0.058	0.067	0.012	0.063	0.118	0.046	0.118
ϵ_{ep_h}	−0.033	0.016	−0.038	0.002	0.005	0.002	−0.063	0.018	−0.032
ϵ_{ep_s}	0.208	0.067	0.131	0.219	0.063	0.219	0.100	0.135	0.224
ϵ_{ep_u}	−0.041	0.074	−0.029	0.004	0.066	0.004	0.029	0.021	0.029
ϵ_{ep_m}	−0.201	0.066	0.046	−0.183	0.110	−0.183	−0.253	0.060	−0.072
ϵ_{mp_k}	0.001	0.012	−0.002	0.005	0.002	0.004	0.000	0.027	0.012
ϵ_{mp_h}	0.003	0.004	0.003	0.005	0.003	0.005	0.006	0.025	0.006
ϵ_{mp_s}	−0.001	0.010	−0.001	−0.004	0.003	−0.004	0.032	0.014	0.032
ϵ_{mp_u}	0.022	0.007	0.019	−0.012	0.004	−0.012	0.040	0.020	0.039
ϵ_{mp_e}	−0.023	0.008	0.002	−0.007	0.004	−0.008	−0.037	0.011	−0.002

(1) Median value of the elasticities over the three subsamples, evaluated at the 1985 data and estimated standard error (s.e.).

Table D3: Output and time elasticities, pooled data

	Box-Cox			Normalized quadratic			Translog		
	unrestricted		restricted	unrestricted		restricted	unrestricted		restricted
	median	s.e.	median	median	s.e.	median	median	s.e.	median
Output elasticities									
ϵ_{cy}	0.873	0.014	0.847	0.798	0.014	0.798	0.890	0.012	0.883
ϵ_{ky}	0.426	0.039	0.400	0.203	0.033	0.199	0.468	0.047	0.459
ϵ_{hy}	0.503	0.151	0.470	0.518	0.133	0.530	0.673	0.141	0.611
ϵ_{sy}	0.631	0.027	0.613	0.517	0.025	0.517	0.650	0.030	0.652
ϵ_{uy}	0.748	0.045	0.713	0.366	0.049	0.368	0.893	0.065	0.889
ϵ_{ey}	0.694	0.100	0.693	0.669	0.146	0.662	0.634	0.114	0.553
ϵ_{my}	1.031	0.034	1.011	1.045	0.018	1.045	1.020	0.015	1.018
Impact of time									
ϵ_{ct}	−0.004	0.000	−0.004	−0.003	0.000	−0.003	−0.005	0.000	−0.005
ϵ_{kt}	−0.005	0.001	−0.006	−0.003	0.001	−0.003	−0.006	0.002	−0.009
ϵ_{ht}	0.024	0.006	0.025	0.009	0.003	0.011	0.027	0.007	0.033
ϵ_{st}	−0.006	0.001	−0.006	−0.003	0.000	−0.003	−0.006	0.001	−0.005
ϵ_{ut}	−0.047	0.002	−0.046	−0.024	0.002	−0.024	−0.051	0.002	−0.051
ϵ_{et}	−0.025	0.004	−0.025	−0.012	0.002	−0.013	−0.032	0.005	−0.031
ϵ_{mt}	0.004	0.001	0.004	0.000	0.000	0.000	0.004	0.001	0.004

(1) Median value of the elasticities evaluated at the 1985 data and estimated standard error (s.e.).

Table D4: Output and time elasticities, sample split (a)

	Box-Cox			Normalized quadratic			Translog		
	unrestricted		restricted	unrestricted		restricted	unrestricted		restricted
	median	s.e.	median	median	s.e.	median	median	s.e.	median
Output elasticities									
ϵ_{cy}	0.854	0.021	0.844	0.803	0.021	0.803	0.889	0.022	0.883
ϵ_{ky}	0.547	0.046	0.480	0.399	0.099	0.399	0.639	0.051	0.667
ϵ_{hy}	0.942	0.110	0.905	1.013	0.111	1.013	0.902	0.180	0.902
ϵ_{sy}	0.662	0.056	0.647	0.579	0.037	0.580	0.746	0.061	0.729
ϵ_{uy}	0.715	0.040	0.715	0.331	0.100	0.332	0.759	0.048	0.759
ϵ_{ey}	0.660	0.044	0.674	0.502	0.080	0.502	0.649	0.084	0.579
ϵ_{my}	0.985	0.022	0.990	1.035	0.024	1.035	0.963	0.022	0.963
Impact of time									
ϵ_{ct}	−0.002	0.001	−0.002	−0.002	0.001	−0.002	−0.003	0.001	−0.003
ϵ_{kt}	−0.002	0.001	−0.006	0.000	0.001	−0.001	−0.006	0.002	−0.007
ϵ_{ht}	0.029	0.002	0.028	0.019	0.001	0.018	0.030	0.007	0.030
ϵ_{st}	−0.004	0.002	−0.004	−0.001	0.001	−0.001	−0.003	0.001	−0.002
ϵ_{ut}	−0.043	0.002	−0.043	−0.024	0.003	−0.024	−0.044	0.002	−0.044
ϵ_{et}	−0.023	0.002	−0.023	−0.017	0.004	−0.017	−0.027	0.004	−0.027
ϵ_{mt}	0.005	0.001	0.005	0.002	0.001	0.002	0.004	0.001	0.004

(1) Median value of the elasticities over the three subsamples, evaluated at the 1985 data and estimated standard error (s.e.).

References

Baltagi, B.H., 1995, *Econometric Analysis of Panel Data*, Wiley.

Baltagi, B.H., J.M. Griffin, and W. Xiong, 2000, 'To Pool or Not to Pool: Homogeneous Versus Heterogeneous Estimators Applied to Cigarette Demand', *Review of Economics and Statistics*, 82, pp. 117-126.

Bergström V. and E. E. Panas, 1992, 'How Robust is the Capital-Skill Complementarity Hypothesis?', *Review of Economics and Statistics,* 74, pp.540-552.

Berman E., J. Bound and Z. Griliches, 1994, 'Changes in the Demand for Skilled Labor within US Manufacturing, Evidence from the Annual Survey of Manufactures', *Quarterly Journal of Economics,* 109, pp.367-397.

Berndt, E. R., and Khaled, M. S., 1979, 'Parametric Productivity Measurement and Choice among Flexible Functional Forms', *Journal of Political Economy,* 87, pp.1220-1245.

Betts, J.R., 1997, 'The Skill Bias of Technological Change in Canadian Manufacturing Industries', *Review of Economics and Statistics,* 79, pp.146-150.

Dhrymes, P.J., 1994, 'Specification Tests in Simultaneous Equation Systems', *Journal of Econometrics*, 64, pp.45-76.

Diewert, W.E., and T.J. Wales, 1987, 'Flexible Functional Forms and Global Curvature Conditions', *Econometrica,* 55, pp.47-68.

Diewert, W. E., and T. J. Wales, 1992, 'Quadratic Spline Models for Producer's Supply and Demand Functions', *International Economic Review*, 33, 705-722.

Falk, M., and B.M. Koebel, 2000, 'Outsourcing of Services, Imported Materials and the Demand for Heterogeneous Labour: An Application of a Generalised Box-Cox Cost Function', Discussion Paper No. 00-51, ZEW, Mannheim.

Fitzenberger, B., 1999, *Wages and employment across skill groups: An analysis for West Germany*, ZEW Economic Studies 6, Heidelberg, Physica.

Gagné, R., and P. Ouellette, 1998, 'On the Choice of Functional Forms: Summary of a Monte Carlo Experiment', *Journal of Business & Economic Statistics*, 16, pp.118-124.

Gouriéroux, C., and A. Monfort, 1989, *Statistique et Modèles Économétriques*, 2 Volumes, Economica, Paris.

Gouriéroux, C., A. Holly, and A. Monfort, 1982, 'Likelihood Ratio Test, Wald, and Kuhn-Tucker Test in Linear Models with Inequality Constraints on the Regression Parameters', *Econometrica*, 50, pp.63-80.

Gouriéroux, C., A. Monfort, and A. Trognon, 1985, 'Moindres Carrés Asymptotiques', *Annales de l'INSEE*, 58, pp.91-122.

Härdle, W., and J.D. Hart, 1992, 'A Bootstrap Test for Positive Definiteness of Income Effect Matrices', *Econometric Theory*, 8, pp276-290.

Härdle, W., W. Hildenbrand and M. Jerison, 1991,'Empirical Evidence on the Law of Demand', *Econometrica*, 59, pp.1525-1549.

Hamermesh, D.S., 1993, *Labor Demand*, Princeton: Princeton University Press.

Janz, N. and B.M. Koebel, 2000, 'Capital Formation in Factor Demand Models: An Empirical Evaluation of Some Common Assumptions', *mimeo.*

Kodde, D.A., and F.C. Palm, 1986, 'Wald Criteria for Jointly Testing Equality and Inequality Restrictions', *Econometrica*, 54, pp.1243-1248.

Kodde D.A. and F.C. Palm, 1987, 'A Parametric Test of the Negativity of the Substitution Matrix', *Journal of Applied Econometrics,* 2, pp. 227-235.

Kodde, D. A., F.C. Palm and G.A. Pfann, 1990, 'Asymptotic Least-Squares Estimation Efficiency Considerations and Estimation', *Journal of Applied Econometrics*, 5, pp.229-243.

Koebel, B.M., 1998, 'Tests of Representative Firm Models: Results for German Manufacturing Industries', *Journal of Productivity Analysis*, 10, pp.251-270.

Koebel, B.M., 2002, 'Can Aggregation Across Goods be Achieved by Neglecting the Problem? — Property Inheritance and Aggregation Bias', *International Economic Review*, (forthcoming).

Lansink, A.O., and G. Thijssen, 1998, 'Testing among Functional Forms: An Extension of the Generalized Box-Cox Fornulation', *Applied Economics*, 30, pp.1001-1010.

Lau, L.J., 1978, 'Testing and Imposing Monoticity, Convexity and Quasi-Convexity Constraints', in M. Fuss and D. McFadden (eds.) *Production Economics: A dual Approach to Theory and Applications*, Volume 1, North Holland, pp.409-453.

Magnus, J.R., 1985, 'On Differentiating Eigenvalues and Eigenvectors', *Econometric Theory,* 1, pp.179-191.

Moschini, G., 1999, 'Imposing Local Curvature Conditions in Flexible Demand Systems', *Journal of Business & Economic Statistics*, 17, pp.487–490.

Ruud, P. A., 1997, 'Restricted Least Squares Subject to Monotonicity and Concavity Constraints', in D.M Kreps and K.F. Wallis, *Advances in economic and econometrics: theory and applications*, Volume III, Cambridge University Press.

Ryan, D.L., and T.J. Wales, 1998, 'A Simple Method for Imposing Local Curvature in Some Flexible Consumer-Demand Systems', *Journal of Business & Economic Statistics*, 16, pp.331–337.

Ryan, D.L., and T.J. Wales, 2000, 'Imposing Local Concavity in the Translog and Generalized Leontief Cost Functions', *Economic Letters*, 67, pp.253–260.

Thompson, P., and T.G. Taylor, 1995, 'The Capital-Energy Substitutability Debate: A new look', *Review of Economics and Statistics,* 77, pp.565-569.

Wolak, F.A., 1989,'Local and Global Testing of Linear and Nonlinear Inequality Constrainst in Nonlinear Econometric Models', *Econometric Theory*, 5, pp.1-35.

Wooldridge, J.M., 1992, 'Some Alternatives to the Box-Cox Regression Model', *International Economic Review*, 33, pp.935-955.