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

Bertrand KOEBEL
Otto-von-Guericke University, D39016 Magdeburg, Germany (*koebel@ww.uni-magdeburg.de*)

Martin FALK
Centre for European Economic Research (ZEW), D68034 Mannheim, Germany (*falk@zew.de*)

François LAISNEY
BETA-THEME Louis Pasteur University, F67000 Strasbourg, France (*laisney@zew.de*)

We present a new method for imposing and testing concavity of cost functions using asymptotic least squares, which can be easily implemented even for nonlinear cost functions. We provide an illustration for a (generalized) Box–Cox cost function with six inputs: capital, labor disaggregated in three skill levels, energy, and intermediate materials. We present a parametric concavity test and compare price elasticities when curvature conditions are imposed versus when they are not. Although concavity is statistically rejected, estimates are not very sensitive to its imposition. We find stronger substitution between the different type of labor than between any other two inputs.

KEY WORDS: Asymptotic least squares; Concavity; Inequality restriction.

1. INTRODUCTION

Most empirical studies in production analysis are based on functional forms that must satisfy some curvature conditions to be compatible with microeconomic theory. The aim of this article 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, but it is increasingly common to directly impose concavity (locally or globally) on the parameters. This can be achieved by restricting the eigenvalues of the Hessian matrix with respect to prices as Gallant and Golub (1984) do, or by relying on the Cholesky decomposition of the Hessian as shown by Lau (1978) and realized by Diewert and Wales (1987). A Bayesian method for imposing concavity has been proposed by Terrell (1996) for three models that are linear in parameters. This method's extension to nonlinear models, although theoretically unproblematic, is probably cumbersome. More recently, Ryan and Wales (1998, 2000) and Moschini (1999) discussed further techniques to impose concavity.

Nonetheless, few contributions formally test whether concavity is statistically rejected by the data. The work of 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 stand point, and 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, e.g., Koebel 2002 on this point). In this context, the a priori imposition of concavity may lead to estimation bias.

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

The solution that we propose makes use of a minimum distance or asymptotic least squares estimator, proposed by Gouriéroux, Monfort, and Trognon (1985) and Kodde, Palm, and Pfann (1990). Concavity is imposed in two stages. First, we estimate the unrestricted parameters to obtain estimate $\widehat{\underline{H}}$; this typically will not be negative semidefinite. Second, the difference between H^0 and $\widehat{\underline{H}}$ is minimized (for an appropriate metric) to obtain the concavity restricted estimates $\widehat{\underline{H}}^0$. We then present a parametric test for the concavity of the cost

Journal of Business & Economic Statistics
April 2003, Vol. 21, No. 2
DOI 10.1198/073500103288618981

function in prices. The method that we rely on for imposing concavity can be used simultaneously for testing this assumption, by testing whether the matrix $\widehat{H}^0 - \widehat{H}$ is statistically different from 0.

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 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 (e.g., by means of wage subsidies). This information cannot be recovered from models considering aggregate labor. In this article, 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 to explain the evolution of different input demands.

Because concavity rejection may in fact be attributable to inappropriate specification of the functional form, we retain a generalized BC formulation that nests several usual models. Although the generalized BC is nonlinear in parameters, concavity is imposed, and the parameters are estimated without great difficulties. We test concavity for several specifications, and investigate whether certain functional forms are more likely to fulfill concavity than others. Furthermore, we compare elasticities when curvature conditions are imposed and when they are not.

The factor demand system is estimated for 31 German manufacturing industries for the period 1978–1990. 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 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 in Figure 1, which describes the evolution of the share of each skill in aggregate labor, with h_t, s_t, and u_t 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 technologic 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.

Sections 2 and 3 are devoted to the techniques used to impose and to test concavity. The generalized BC 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 elasticities are discussed in Section 7. Some conclusions are presented in Section 8.

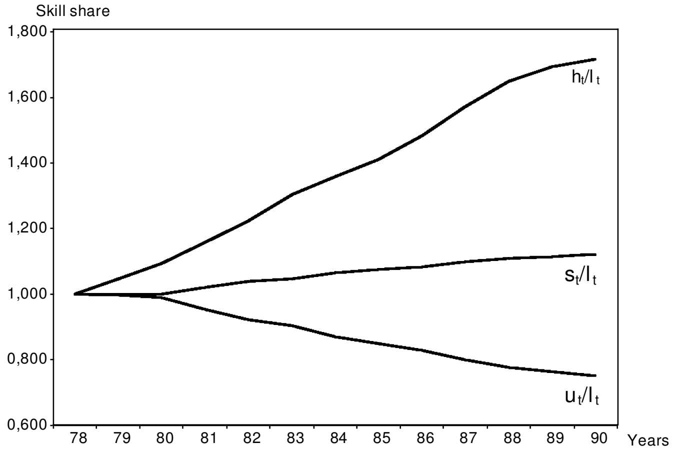

Figure 1. Evolution of the Shares of Three Types of Qualification (h = high skill, s = skill, u = unskilled) in Aggregate Manufacturing Employment (l = labor) in West Germany; Basis 1978 = 1.00.

2. PARAMETER ESTIMATION UNDER CONCAVITY RESTRICTION

The technology 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 technology parameters. The cost function c gives the minimal value in x of the product $p'x$ that can be achieved for given prices p and the technology constraint; that is,

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

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 is obtained by applying Shephard's lemma to c,

$$x^*(p, z, \alpha) = \frac{\partial c(p, z; \alpha)}{\partial p} \geq 0. \tag{2}$$

This last inequality means that input price increases leads to higher costs; that is, c is monotone (nondecreasing) in p.

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,

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

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

$$\frac{\partial^2 c(p, z; \alpha)}{\partial p \partial p'} p = 0, \tag{4}$$

and 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: $p=\bar{p}$ and $z=\bar{z}$.

Let N denote the number of observations. 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 error terms. We assume that Ω_ε, the conditional variance of ε, is consistently estimated by $\widehat{\Omega}_\varepsilon$. The unrestricted least squares estimator $\hat{\alpha}$ is defined as

$$\hat{\alpha} = \arg\min_{\alpha}(X - X^*(\alpha))'\widehat{\Omega}_\varepsilon^{-1}(X - X^*(\alpha)). \tag{5}$$

The concavity-restricted least squares estimator is obtained as

$$\hat{\alpha}^0 = \arg\min_{\alpha}\{(X - X^*(\alpha))'\widehat{\Omega}_\varepsilon^{-1} \times (X - X^*(\alpha)) : v'\overline{H}v \le 0, \forall v \in \mathbb{R}^{S_p}\}. \tag{6}$$

Because the matrix H in (3) in general depends not solely on α, but also on price and output levels, concavity is only imposed locally in (6) (at $p=\bar{p}$ and $z=\bar{z}$) and will not necessarily be fulfilled at other observation points. Hence the notation $\overline{H}$ in (6). The method that we propose could, however, be easily 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 that is attractive because of its simplicity and that 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) showed that restricting the parameters α to fulfill negative semidefiniteness 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 semidefiniteness of $\overline{H}$ for any value of the remaining parameters α_B. Note that for those functional forms that 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. 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 $\overline{A}$ depend only on the concavity driving parameters $\alpha_A = \text{vecli}\overline{A}$, and $\overline{B}$ depends only on α_B and not on α_A. The operator vecli 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 ambiguity of the definition is only superficial. We can therefore write

$$H(\bar{p},\bar{z};\alpha_A,\alpha_B) = A(\bar{p},\bar{z};\alpha_A) + B(\bar{p},\bar{z};\alpha_B) \tag{7}$$

or, more succinctly, $\overline{H} = \overline{A}+\overline{B}$. Negative semidefiniteness 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 semidefinite. For this purpose, Ryan and Wales (1998) proposed reparameterizing 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 estimating the parameters of U and $\overline{B}$ instead of $\overline{A}$ and $\overline{B}$. Negative semidefiniteness of $\overline{H}$ is then achieved by construction. Because $\overline{A} = -U'U - \overline{B}$, 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 $\eta_H^0(u) = \text{vecli}H^0$ 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 $u = \text{vecli}U$. Instead of estimating the parameters $\alpha = (\alpha_A', \alpha_B')'$ of the cost function, Ryan and Wales estimated $(u', \alpha_B')'$ by solving a nonlinear least squares problem of the type

$$\min_{u,\alpha_B}(X - X^*(u,\alpha_B))'\widehat{\Omega}_\varepsilon^{-1}(X - X^*(u,\alpha_B)). \tag{8}$$

Identification of the parameter vector $\alpha = (\alpha_A', \alpha_B')$, which is of interest for the computation of elasticities, may then be obtained from (7).

This approach presents three main drawbacks. First, obtaining convergence may be difficult. By implementing (8) in turn for 29 industrial sectors, Koebel (1998) encountered 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 (7) is not possible for every functional form, and 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; Moschini 1999 on this last point). Third, tests for the concavity assumption are not provided.

Instead of relying on (8), we could estimate the concavity-restricted parameters via the asymptotically equivalent minimum distance estimator obtained as the solution of

$$\tilde{\alpha}^0 = \arg\min_{\alpha}\{(\hat{\alpha}-\alpha)'\widehat{\Omega}_\alpha^{-1}(\hat{\alpha}-\alpha) : v'\overline{H}v \le 0, \forall v \in \mathbb{R}^{S_p}\}, \tag{9}$$

where $\hat{\alpha}$ denotes the unrestricted estimate of α and $\widehat{\Omega}_\alpha$ is a consistent estimate of the variance matrix of $\hat{\alpha}$. In (9), 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 (6) and (9) has been discussed by Gouriéroux and Monfort (1989, chap. XXI).

In general, the inequality constraint $v'\overline{H}v \le 0$ in (9) cannot be explicitly imposed on the parameters α; 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 justified in Proposition 1. In the first stage, the parameters $\hat{\eta}_H^0$ of the concavity-restricted Hessian matrix $\widehat{\overline{H}}^0$ are determined as the solution of

$$\min_{u}(\hat{\eta}_H - \eta_H^0(u))'\widehat{\Omega}_H^{-1}(\hat{\eta}_H - \eta_H^0(u)) \equiv d, \tag{10}$$

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

$$\widehat{\Omega}_H \equiv \frac{\partial g}{\partial \alpha'}(\hat{\alpha})\widehat{\Omega}_\alpha \frac{\partial g'}{\partial \alpha}(\hat{\alpha}). \tag{11}$$

Gouriéroux and Monfort (1989) and Wolak (1989) showed that the minimum achieved in (9) is asymptotically equivalent to the Wald statistic d in (10).

From (10), we obtain $\hat{\eta}_H^0$, but in most cases these estimates (in number S_{η_H}) do not enable identification of the S_α concavity-restricted parameters $\hat{\alpha}^0$ of the cost function. Therefore, a second stage is needed to identify the parameters of interest $\hat{\alpha}^0$. Identification can be achieved by adapting the asymptotic least squares framework proposed by Gouriéroux et al. (1985). The assertions and propositions that follow are proven in Appendix A. The relationship between the S_{η_H} restricted parameters $\hat{\eta}_H^0$ and the S_α structural parameters $\hat{\alpha}^0$ can be written as

$$\hat{\alpha}^0 = \arg\min_\alpha \{(\hat{\alpha} - \alpha)'\widehat{\Omega}_\alpha^{-1}(\hat{\alpha} - \alpha) : \hat{\eta}_H^0 = g(\alpha)\}, \tag{12}$$

and the solution to this problem is asymptotically equivalent to

$$\hat{\alpha}^0 = \hat{\alpha} + \widehat{\Omega}_\alpha \frac{\partial g'}{\partial \alpha}(\hat{\alpha}) \left(\frac{\partial g}{\partial \alpha'}(\hat{\alpha})\widehat{\Omega}_\alpha \frac{\partial g'}{\partial \alpha}(\hat{\alpha}) \right)^{-1} (\hat{\eta}_H^0 - g(\hat{\alpha})). \tag{13}$$

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

Proposition 1. Under the assumption that each of the problems (9), (10), and (12) has a unique solution, the following properties are verified under the null of concavity:

a. The solution $\tilde{\alpha}^0$ of (9) and the solution $\hat{\alpha}^0$ of (12) are asymptotically equivalent.

b. The solution $\hat{\eta}_H^0$ of (10) is asymptotically equivalent to $g(\tilde{\alpha}^0)$ and to $g(\hat{\alpha}^0)$.

c. The minimum distances achieved in problem (9), (10), and (12) are asymptotically equivalent.

d. An asymptotic solution of (12) is given by (13).

Part a of Proposition 1 justifies our two-step procedure for solving (9). Part b allows us to retain the statistic $\hat{\eta}_H^0$ obtained by solving (10) as an estimator for $g(\alpha^0)$. Part c provides a rationale for computing the Wald-type test for the null of concavity using the minimum value achieved in (10). Part d justifies using (13) for determination of the concavity-restricted parameters.

Concerning the asymptotic distribution of the estimators $\hat{\eta}_H^0$ and $\hat{\alpha}^0$ under H_0, we must distinguish the case where the true values saturate some or all constraints. If they do not, then these asymptotic distributions are $N(\alpha, \Omega_\alpha)$ and $N(\eta_H, \Omega_H)$. If some constraints are saturated by the true value, however, then the distributions of $\hat{\eta}_H^0$ and $\hat{\alpha}^0$ become quite complex (see Gouriéroux, Holly, and Monfort 1982; Kodde and Palm 1986; Wolak 1989).

A special case of this minimum distance estimator of $\hat{\alpha}^0$ was used by Koebel (1998) for estimation of the concavity-restricted parameters of a normalized quadratic cost function. For this functional form, the matrix $B(\bar{p}, \bar{z}; \alpha_B)$ vanishes in the expression (7), and the concavity-restricted parameters α_A^0 can be estimated and identified in one stage. The main advantage of using (10) and (13) rather than (8) is that convergence is obtained much more easily. As is shown in the next section, (10) is also useful for testing the validity of the concavity restrictions.

Note that the monotonicity property (2) can be imposed locally using a similar technique. For instance, let us define $\hat{\zeta} \equiv x^*(\bar{p}, \bar{z}; \hat{\alpha})$ and $\zeta^0(v) \equiv (v_1^2, \ldots, v_{S_x}^2)'$, where the v_j are parameters to be estimated. Then it is possible in a first stage to minimize the distance between $\hat{\zeta}$ and $\zeta^0(v)$ with respect to v to obtain the monotonocity-restricted estimates $\hat{\zeta}^0$. The monotonicity-restricted parameters can then be obtained in a second stage in a similar setup as the one presented earlier. It is of course also possible to impose monotonicity and concavity simultaneously along the same lines.

3. TESTING CONCAVITY

Tests of the definiteness of a matrix have been presented in the context of demand analysis. For example, Härdle and Hart (1992) and Härdle et al. (1991) tested whether the highest nonzero eigenvalue of $\widehat{\widehat{H}}$ is significantly negative. Kodde and Palm (1987) preferred to consider all eigenvalues simultaneously and propose a distance test based on

$$d_{KP} \equiv \min_{\lambda \le 0} (\hat{\lambda} - \lambda)'\widehat{\Sigma}^-(\hat{\lambda} - \lambda), \tag{14}$$

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

$$\widehat{\Sigma} = \frac{\partial \lambda}{\partial\, \text{vec}'\overline{H}}(\widehat{\widehat{H}})\widehat{\mathbf{\Omega}}_H \frac{\partial \lambda'}{\partial\, \text{vec}\overline{H}}(\widehat{\widehat{H}}),$$

where $\widehat{\mathbf{\Omega}}_H$ is the variance matrix of $\text{vec}\widehat{\widehat{H}}$. Note that $\widehat{\mathbf{\Omega}}_H$ is different from $\widehat{\Omega}_H$, which is the variance matrix of $\text{vecli}\widehat{\widehat{H}}$. Whereas the latter has full rank, the former is singular. A generalized inverse of $\widehat{\Sigma}$ must be considered in (14), because 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 (1987).

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

$$d_{KP} \ge d,$$

$$d_{KP} \overset{a}{=} d.$$

Briefly, Proposition 2 states that in small samples $d_{KP} \ge d$, 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 $\widehat{\bar{H}}$ and the negative semidefinite matrix $\widehat{\bar{H}}^0$. The distance d may, however, be more useful than d_{KP} for three reasons. First, computation of d is somewhat simpler, because we do not have to calculate the matrix of derivatives of the eigenvalues with respect to the parameters $\partial\lambda/\partial \text{vec}'\bar{H}$. Second, the statistic d can be computed even when the 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 $\hat{\alpha}^0_A$ by solving (10); this is not the case when using (14).

Does the assumption that the eigenvalues are differentiable strongly restrict the applicability of Proposition 2? The following result shows that the set of matrices that have multiple eigenvalues is of measure 0, and therefore the assumption that the eigenvalues are differentiable does not much restrict the applicability of Proposition 2.

Proposition 3. a. The eigenvalue $\bar{\lambda}_1$ is differentiable with respect to η_H if and only if $\bar{\lambda}_1$ is simple.
b. The set of all matrices $\bar{H}$ that have multiple eigenvalues is of Lebesgue measure 0.

Proposition 3 means that almost all matrices have differentiable eigenvalues. Hence the technical problem related to the nondifferentiability of some eigenvalues that may arise when Proposition 2 is applied occurs for only a small set of matrices with measure 0.

Gouriéroux et al. (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-square distributions,

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

where the weight $\bar{w}_j$ denotes the probability that j of the $S_p - 1$ eigenvalues of $\bar{H}$ are negative. Because 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

To avoid imposing a priori restrictions on an unknown technological structure, researchers have relied on flexible functional forms that can be interpreted as (local) second-order approximations to an arbitrary cost function. Translog (TL), generalized Leontief (GL), and normalized quadratic (NQ) cost functions have often been used for estimating price elasticities. We consider a generalized BC cost function that nests the former usual specifications. In contrast to the BC formulations of Berndt and Khaled (1979) and Lansink and Thijssen (1998), our specification nests both the NQ and the TL cost functions.

We apply the BC transformation to the explanatory variables p_{it} and z_{it}; for $\gamma_1 \neq 0$, let

$$Z_{jit} = \frac{z_{jit}^{\gamma_1} - 1}{\gamma_1}, \qquad j = 1, \ldots, S_z,$$

and

$$P_{jit} = \frac{(p_{jit}/p'_{it}\theta_i)^{\gamma_1} - 1}{\gamma_1}, \qquad j = 1, \ldots, S_p.$$

For $\gamma_1 = 0$, let $Z_{jit} = \ln z_{jit}$ and $P_{jit} = \ln(p_{jit}/p'_{it}\theta_i)$. The term $p'_{it}\theta_i$ 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 $\bar{x}_i/\bar{c}_i$, where $\bar{x}_i$ and $\bar{c}_i$ denote some average inputs and costs defined later. The choice of a kind of Laspeyres cost index $p'_{it}\theta_i$ for normalization is appealing because it is independent of the units of measurement of prices and of quantities (i.e., satisfies the index theoretical dimensionality and commensurability axioms).

The specification of the cost function is

$$c^*(p_{it}, z_{it}; \alpha_i) = p'_{it}\bar{x}_i(\gamma_2 C^*(p_{it}, z_{it}; \beta_i, \gamma_1) + 1)^{1/\gamma_2}, \qquad (15)$$

for $\gamma_2 \neq 0$ and $c^* = p'_{it}\bar{x}_i \exp(C^*)$ for $\gamma_2 = 0$, where

$$\begin{aligned} C^* &= C(P_{it}, Z_{it}; \beta_i) \\ &= \beta_{0i} + (P'_{it}, Z'_{it})B_{1i} + \frac{1}{2}(P'_{it}, Z'_{it})B_2\begin{pmatrix} P_{it} \\ Z_{it} \end{pmatrix} \\ &= \beta_{0i} + (P'_{it}, Z'_{it})\begin{pmatrix} B_{pi} \\ B_z \end{pmatrix} + \frac{1}{2}(P'_{it}, Z'_{it})\begin{pmatrix} B_{pp} & B_{pz} \\ B_{zp} & B_{zz} \end{pmatrix}\begin{pmatrix} P_{it} \\ Z_{it} \end{pmatrix}. \end{aligned} \qquad (16)$$

In c^*, the technology parameters to be estimated are gathered in the vector $\alpha_i = (\beta'_i, \gamma_1, \gamma_2)'$. The matrices B_{1i} and B_2 contain the parameters of β_i and are of size $(S_p + S_z) \times 1$, and $(S_p + S_z) \times (S_p + S_z)$. It can be directly seen that the cost function c^* is linearly homogeneous in prices. The term $p'_{it}\bar{x}_i$ 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 BC 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 (15), 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 $p'_{it}\bar{x}_i$ 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 with the number of parameters entailed by the specification (16), 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 0 in p_{it} imply $(S_p + S_z)(S_p + S_z - 1)/2$ and $1 + S_p + S_z$ additional restrictions on C^*. Hence, for C^* to be a flexible function, it is only necessary that it 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

$$\begin{aligned} &B_{pp} = B'_{pp}, \qquad B_{zz} = B'_{zz}, \qquad B_{pz} = B'_{zp}, \\ &\iota'_{S_p}B_{pi} = 1, \qquad \iota'_{S_p}B_{pp} = 0, \qquad \iota'_{S_p}B_{pz} = 0, \end{aligned} \qquad (17)$$

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

From (15), (16), and (17), 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 NQ cost function is obtained. The 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 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 foregoing BC specification is a flexible functional form; because (15) entails several flexible functional forms as special cases, the BC cost function itself is flexible.

The system of input demands $x^*(p_{it}, z_{it}; \alpha_i)$ 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

$$x^*_{it} = \bar{x}_i c^*_{it}/(p'_{it}\bar{x}_i) + p'_{it}\bar{x}_i(c^*_{it}/(p'_{it}\bar{x}_i))^{(1-\gamma_2)}\frac{\partial P'_{it}}{\partial p_{it}}\frac{\partial C_{it}}{\partial P_{it}}, \quad (18)$$

where

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

and

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

By convention, $\hat{p} \equiv \operatorname{diag}(p_{it})$ 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 0 in p_{it}. Hence, for the specification (18), the adding-up condition $p'_{it}x^* = c^*$ is automatically satisfied.

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

$$\begin{aligned}\frac{\partial^2 c^*}{\partial p_{it}\partial p'_{it}} &= \bar{x}_i(c^*_{it}/p'_{it}\bar{x}_i)^{(1-\gamma_2)}\frac{\partial C_{it}}{\partial P'_{it}}\frac{\partial P_{it}}{\partial p'_{it}} \\ &+ (c^*_{it}/p'_{it}\bar{x}_i)^{(1-\gamma_2)}\frac{\partial P'_{it}}{\partial p_{it}}\frac{\partial C_{it}}{\partial P_{it}}\bar{x}'_i \\ &+ p'_{it}\bar{x}_i(1-\gamma_2)(c^*_{it}/p'_{it}\bar{x}_i)^{(1-2\gamma_2)} \\ &\times \left(\frac{\partial P'_{it}}{\partial p_{it}}\frac{\partial C_{it}}{\partial P_{it}}\right)\left(\frac{\partial C_{it}}{\partial P'_{it}}\frac{\partial P_{it}}{\partial p'_{it}}\right) + p'_{it}\bar{x}_i(c^*_{it}/p'_{it}\theta_i)^{(1-\gamma_2)} \\ &\times \left[\frac{\partial P'_{it}}{\partial p_{it}}B_{pp}\frac{\partial P_{it}}{\partial p'_{it}} + \left(\frac{\partial C_{it}}{\partial P'_{it}}\frac{\partial^2 P_{it}}{\partial p_{jt}\partial p_{ht}}\right)\right]. \quad (19)\end{aligned}$$

Evaluating this matrix at $(\bar{p}_{it}, \bar{z}_{it})$, we see that it does not admit an additively separable representation such as (7). For this reason, it is more convenient to apply the method presented in Section 2 for determination of the concavity-constrained estimates. Evaluating (19) at the unrestricted parameter values $\hat{\alpha}_i$ yields $\hat{\eta}_H$, which can in turn be replaced in problem (10) to derive the minimum distance estimates $\hat{\eta}^0_H$ for the parameters of the restricted Hessian matrix $\nabla^2_{pp}c\left(\bar{p}, \bar{z}; \hat{\alpha}^0\right)$. Knowing the value of $\hat{\eta}^0_H$, the restricted parameters $\hat{\alpha}^0$ of interest for the computation of the different elasticities can be computed using (13).

5. EMPIRICAL IMPLEMENTATION

We first briefly describe the dataset that we use, then present some preliminary results aimed to precise the sample split and the specification on which we rely for testing concavity.

5.1 Data Description

Given the available data, 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} denotes low-skilled or unskilled labor. 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 industry. Other explanatory variables entering the cost function are the level of production, y_{it}, and a time trend, t. These variables are regrouped in a vector $z_{it} = (y_{it}, t)'$. The BC transformation is not applied to t; hence $Z_{it} = ((y^{\gamma_1}_{it} - 1)/\gamma_1, t)$ in (16). Total costs of production are defined by $c_{it} = p'_{it}x_{it}$.

The data used consist of a panel of 31 out of 32 German two-digit manufacturing industries observed over the period 1978–1990. One industry (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 available only from 1978 onward. Because of data-related difficulties appearing with the German reunification, we prefer not to use post reunification data. Most of our data are 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 yearly information on employment by skill category and by industry as of June 30, for all employees paying Social Security contributions. Labor is split into three groups: group 1 (high-skilled) is defined as workers with a university or polytechnical degree, group 2 (skilled) comprises those having completed vocational training as well as technicians and foremen, and group 3 (unskilled) comprises workers without formal qualifications. From this dataset, we calculate the shares of the three skill groups in employment and multiply these by total employment, available for each industry from the national accounts, to obtain h_{it}, s_{it}, and u_{it}.

Information on earnings is taken from the IAB_S dataset for skilled and unskilled labor and from the Federal Statistical Office (Löhne und Gehälter Statistik) for high-skilled labor. 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. More information on the data is provided in the discussion paper version (Koebel, Falk, and Laisney 2000).

To reduce heteroscedasticity, input demands are divided by the output level,

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

For the 1978–1990 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 0 mean and a constant variance matrix Ω_ν and that it is uncorrelated with the regressors. (Note that the matrix Ω_ν is not singular, as would be the case with a system of shares.) We thus obtain maximum likelihood estimates under the assumption $\nu \sim N(0, \Omega_\nu)$.

In defining the parameters $\bar{x}_i$ and θ_i, we try to avoid any correlation with the error term ν_{it}. The following choice seems to be convenient for that purpose:

$$\theta_i = \frac{\bar{x}_i}{\bar{c}_i} = \frac{\frac{1}{N-1}\sum_{n\neq i} x_{n,1}}{\frac{1}{N-1}\sum_{n\neq i} c_{n,1}}.$$

In this definition, industry i has been excluded in the summations to ensure that θ_i and $\bar{x}_i$ are uncorrelated with ν_{it} (as ν_{it} and ν_{jt} are uncorrelated for $i \neq j$). Thus $p'_{it}\theta_i$ corresponds effectively to a kind of Laspeyres price index for total costs, with the basis year chosen to be $t = 1$. The empirical results are robust to the exclusion of the basis year in the regression.

5.2 Preliminary Results

First, the parameters α_i have been estimated by assuming that relation (20) is valid for all industries in our sample. For $S_p = 6$, (18) entails 218 free parameters (among which 186 industry dummies), which have to be estimated on the basis of $31 \times 13 \times 6 = 2{,}418$ observations. To account for sectoral differences, the BC specification (16) includes some industry-specific coefficient β_{0i} and B_{pi}, but in fact the remaining coefficients might also differ across industries. Given the relatively short time dimension, the parameters of the BC model cannot be estimated for each industry separately. Therefore, we investigate parameter heterogeneity by estimating model (20) for different subgroups of industries. These groups are formed on the basis of similarities (a) in their production, (b) in their size, (c) in their skill structure of labor, and on whether they are (d) labor- or (e) capital-intensive. Within each group, it is assumed that technologies differ only through β_{0i} and B_{pi}. Across groups, technologies can differ in any of the parameters α_i.

The first sample split distinguishes three groups of industries 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 industries in three groups according to their level of production (at 1984 values). The last three sample splits correspond to the size of some cost shares. In each case, we split the 32 industries into 3 groups, each comprising 10 or 11 industries, according to whether they are located in the lower, middle, or upper third 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. For the groups with 10 (resp. 11) industries, there are 92 (98) parameters [among which 60 (66) are industry dummies], which have to be determined using $10 \times 13 \times 6 = 780$ (858) observations. Table 1 summarizes results of likelihood ratio tests for the null of identical BC technologies across industries. In all cases, the pooled model is rejected. Sample splits (a) and (c) yield the highest log-likelihood values.

Even if statistical tests reject the equality of some parameters, there exist arguments in favor of pooling the data. Baltagi (1995, chap. 4) recommended using a mean squared error (MSE) criterion for assessing the poolability of the data, rather than tests on the equality of parameters. In modeling 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 [. . .] heterogeneities" (p. 125). Because a long time period is necessary for such comparisons, we cannot pursue these lines. Instead, we focus on the pooled model and on the two disaggregate models with the highest likelihood. 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 form and for concavity.

Table 1. BC Estimates for Different Sample Splits

	Subsample (i)	*Subsample (ii)*	*Subsample (iii)*	*LR-test*[(1)]
Split (a)	Consumer goods	Investment goods	Intermediate goods	LR test[(1)]
log-L	3,900	2,537	2,656	2034
Split (b)	Small industries	Medium industries	Large industries	LR test[(1)]
log-L	2,799	2,912	2,941	1149
Split (c)	Low skill intensive	Skill intensive	Highly skill intensive	LR test[(1)]
log-L	3,338	2,951	2,651	1727
Split (d)	Not labor intensive	Labor intensive	Highly labor intensive	LR test[(1)]
log-L	3,168	2,740	2,830	1320
Split (e)	Not capital intensive	Capital intensive	Highly capital intensive	LR test[(1)]
log-L	3,056	2,868	2,882	1459

NOTE: See Tables C-1 and C-2 in Appendix C for the composition of the different subsamples.

(1) The likelihood ratio test is calculated as $2(\sum_{j=1}^{3} \log \ell_j - 8{,}077.18)$, where 8,077.18 is the log-likelihood obtained on the pooled sample and $\log \ell_j$ denotes the log-likelihood obtained for the jth group of the corresponding sample split. Under the null, this test statistic is asymptotically $\chi^2(106)$; there are $3 \times 32 - 32 = 64$ slope parameters and $3 \times 21 - 21 = 42$ independent terms in the covariance matrix. The 5% critical value is 131.

Table 2. Log-Likelihood Values of Alternative Specifications

Specification	Sample		γ_1 (t value)		γ_2 (t value)		Log-likelihood
Box–Cox	Pooled		.345	(15.1)	.195	(12.6)	8,077
	Split (a)	(i)	.739	(23.1)	.411	(22.2)	
		(ii)	.070	(1.6)	.087	(6.5)	9,094
		(iii)	.558	(13.1)	.109	(3.8)	
	Split (c)	(i)	.517	(13.5)	.270	(9.2)	
		(ii)	.730	(17.4)	.206	(8.7)	8,941
		(iii)	.856	(32.9)	.694	(25.6)	
Normalized quadratic	Pooled						7,683
	Split (a)		1		1		8,776
	Split (c)						8,675
Generalized Leontief	Pooled						7,516
	Split (a)		1/2		1		8,535
	Split (c)						8,520
Generalized square root	Pooled						7,055
	Split (a)		1		2		8,124
	Split (c)						8,029
Translog	Pooled						7,982
	Split (a)		$\to 0$		$\to 0$		8,876
	Split (c)						8,738
Lin-log	Pooled						7,166
	Split (a)		1		$\to 0$		8,082
	Split (c)						8,171
Log-lin	Pooled						7,582
	Split (a)		$\to 0$		1		8,666
	Split (c)						8,534

5.3 Tests of Nested Specifications

Several usual specifications of input demands are nested within the BC model and thus can be easily tested against it. Table 2 provides the estimates obtained for $\hat{\gamma}_1$ and $\hat{\gamma}_2$, their t statistics, and the log-likelihood values obtained for the pooled and two disaggregate models. The upper part of Table 2 gives the results for the BC specification. Although they differ between subsamples, in all cases $\hat{\gamma}_1$ and $\hat{\gamma}_2$ are between 0 and 1 and are statistically different from 0 and 1. This suggests that the estimated BC departs in significant ways from common functional forms. Only in subsample (a, ii) the assumption $\gamma_1 = \gamma_2$ cannot be rejected; this corresponds to a BC form similar to the one proposed by Berndt and Khaled (1979).

The log-likelihood values for alternative functional forms nested within the BC are reported in Table 2. Among these, the TL achieves the highest likelihood, followed by the NQ. However, likelihood ratio tests reject the null hypothesis that the alternative functional form describes the technology as well as the BC for all specifications and samples considered.

Because all alternative specifications are rejected, only the BC should be retained in the sequel. For comparing alternative functional forms, both in terms of their elasticities and in their ability to satisfy concavity, we continue to consider the NQ and the TL. This permits us to assess whether concavity violation is due to the choice of a particular functional form or is rejected for the bulk of the specifications.

6. CONCAVITY TESTS

In this section we first present further estimates based on the concavity-unrestricted model discussed earlier. Then, using the method outlined in Section 2, we determine the concavity restricted parameters $\hat{\alpha}^0$ and the test statistic $\hat{d}$ for the null of concavity. The results are presented in Table 3. Columns 3–5 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.

Table 3. Tests of Price Concavity of the Cost Function

			Unrestricted estimates			Restricted estimates		
			Highest	No. of positive	% of	Concavity	% of concavity	
Specification	Sample		eigenvalue	eigenvalues	failures	test $\hat{d}$	Reject	Accept
Box–Cox	Pooled		1.9 (4.8)	1.9 (.6)	100.0	68.4 (56.9)	77.9	10.9
		(i)	.6 (1.3)	1.2 (.7)	92.3	27.5 (21.2)	76.9	12.4
	Split (a)	(ii)	3.8 (4.2)	2.1 (.5)	100.0	18.1 (8.8)	82.1	0
		(iii)	1.9 (3.0)	2.2 (.4)	100.0	15.2 (11.0)	58.1	9.4
		(i)	2.6 (5.9)	1.4 (.7)	100.0	25.2 (32.9)	51.7	16.1
	Split (c)	(ii)	.8 (.9)	1.1 (.3)	100.0	27.0 (16.6)	67.7	3.1
		(iii)	1.0 (1.0)	2.3 (.7)	100.0	13.3 (10.7)	41.5	2.3
Normalized	Pooled		1.3 (.2)	2.0 (.0)	100.0	23.0 (0)	100.0	0
quadratic		(i)	2.5 (.3)	2.0 (.0)	100.0	5.0 (0)	0	0
	Split (a)	(ii)	3.4 (.5)	1.0 (.0)	100.0	31.5 (0)	100.0	0
		(iii)	6.2 (.9)	3.0 (.0)	100.0	17.6 (0)	100.0	0
		(i)	6.9 (1.0)	2.0 (.0)	100.0	42.1 (0)	100.0	0
	Split (c)	(ii)	2.8 (.4)	3.0 (.0)	100.0	18.4 (0)	100.0	0
		(iii)	.9 (.1)	2.0 (.0)	100.0	9.4 (0)	0	0
Translog	Pooled		5.4 (15.2)	2.0 (.8)	97.3	102.1 (82.3)	80.4	8.7
		(i)	13.8 (38.5)	2.1 (.9)	89.9	91.4 (75.4)	84.6	15.4
	Split (a)	(ii)	5.3 (5.8)	2.0 (.3)	100.0	17.6 (6.9)	87.2	0
		(iii)	1.9 (3.4)	1.8 (.7)	100.0	37.7 (37.1)	64.1	17.1
		(i)	4.9 (13.6)	1.4 (.8)	92.3	40.3 (50.4)	65.0	23.8
	Split (c)	(ii)	2.2 (3.0)	1.5 (.5)	100.0	66.7 (59.2)	64.6	3.8
		(iii)	1.4 (1.7)	1.9 (.8)	94.6	70.0 (45.6)	76.9	16.9

NOTE: In columns 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 $d_u = 10.371$. For the NQ, the concavity test should take the same value at every observation. Thus the zero standard deviations reported in columns 4 and 6 are not surprising.

The main conclusion that can be drawn from these unrestricted estimates is that they do not often satisfy concavity. For the NQ, concavity is even globally violated for all samples. It is violated in all observations for the BC, except in subsample (a, i), and for the TL except for these three subsamples and for the pooled sample. Thus, although the BC is more flexible than the TL, it cannot be said to satisfy concavity more easily. No clear pattern emerges for the number of positive eigenvalues, but the largest eigenvalues are almost uniformly lower for the BC than for the TL.

The second part of Table 3 (columns 6–8) reports some results on the statistical significance of concavity violations. Because 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). Thus we compute the test statistic taking in turn each observation as the reference point. The average level of $\hat{d}$ over all reference points of the group and its sample standard deviation are given in columns 6 and 7. The null hypothesis of concavity is rejected when the test statistic $\hat{d}$ is found to be significantly different from 0. The percentage of cases for which $\hat{d}$ was found to be significant (nonsignificant) is reported in columns 7 and 8. Because we use the upper and lower bounds proposed by Kodde and Palm (1986) for testing the inequality restrictions, these percentages do not sum to 100, and it is not possible to reach a conclusion for every value of $\hat{d}$.

The test points out that concavity violation is significant on average. For 19 of 21 models, the number of conclusive rejections exceeds that of conclusive failures to reject. In three cases—the BC on subsample (c, iii) and the NQ on subsamples (a, i) and (c, iii)—many test statistics fall in the inconclusive area (i.e., $d_\ell < \hat{d} < d_u$). We refrained from completing Table 3 with the results of monotonicity tests, because monotonicity failures do not appear to be a problem with our data and specifications, as documented in Section 7.1.

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 provide only weak evidence for a rejection, global concavity would be unambiguously rejected. This is in line with the findings of Terrell (1996).

7. RESTRICTED AND UNRESTRICTED ELASTICITIES

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 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). Because the variations over time are not substantial, all elasticities are evaluated at 1984 values. To save space, we report the median value of each elasticity over the industries and its standard error (s.e.), estimated with the delta method. The concavity restriction is imposed for the year 1984 and for the industry producing the median level of output (no. 42), which defines the point $(\bar{p}, \bar{z})$.

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 in elasticities with respect to the functional form retained: ϵ_{hp_h} ranges between -1.53 for the TL and $-.28$ for the NQ, and ϵ_{ep_e} ranges between $-.10$ for the NQ and .59 for the TL. Note that inputs with a low cost share (e.g., energy and high-skilled labor) have particularly variable own-price elasticities across the specifications. For material inputs, the own-price elasticity is rather stable.

Because concavity imposition is in fact valid only if the cost function is price monotonic, we verify whether this condition is fulfilled. For each functional form (BC, NQ, and TL) and each of the sample splits [pooled and split (a)], there is a total of 31×13 points at which the six input demands must be evaluated. In summary, the (concavity-) unrestricted specifications do not often violate price monotonicity, because among $6 \times 6 \times 31 \times 13 = 14{,}508$ possible cases, this occurs for only 1 case. Monotonicity was never violated at the point $(\bar{p}, \bar{z})$ at which concavity was imposed subsequently. After concavity imposition, price monotonicity fails to hold for two observations. Again we find that price monotonicity is never violated at the point $(\bar{p}, \bar{z})$ at which concavity is imposed; thus it turns

Table 4. Own-Price 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
ϵ_{kp_k}	−.163	.054	−.378	−.070	.018	−.082	−.083	.064	−.367
ϵ_{hp_h}	−1.301	.331	−2.350	−.284	.247	−.777	−1.533	.529	−3.479
ϵ_{sp_s}	−.568	.050	−.621	−.254	.032	−.298	−.499	.069	−.616
ϵ_{up_u}	−.347	.124	−.520	−.091	.100	−.167	−.095	.128	−.421
ϵ_{ep_e}	.497	.086	−.396	−.099	.027	−.106	.592	.115	−.017
ϵ_{mp_m}	−.055	.018	−.092	.013	.011	−.010	−.087	.024	−.124

NOTE: Median value of the elasticities evaluated for the 1984 data and estimated standard error (s.e.).

Table 5. Own-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_k}	−.062	.030	−.114	−.068	.050	−.068	−.021	.053	−.136
ϵ_{hp_h}	−.831	.344	−.831	−.050	.059	−.073	−.967	.289	−.967
ϵ_{sp_s}	−.464	.103	−.443	−.045	.023	−.061	−.406	.064	−.497
ϵ_{up_u}	−.331	.089	−.322	.118	.074	−.007	−.174	.104	−.254
ϵ_{ep_e}	.012	.043	−.269	−.064	.139	−.074	.020	.068	−.251
ϵ_{mp_m}	.016	.056	.016	.021	.004	.021	−.056	.070	−.060

NOTE: Median value of the elasticities over the three subsamples, evaluated for the 1984 data and estimated standard error (s.e.).

out not to be necessary to impose price monotonicity in addition to concavity at $(\bar{p}, \bar{z})$. The cost function is also found to be always increasing in output, whichever specification or sample split is retained.

When concavity is locally imposed on the pooled estimates, all own-price elasticities become negative. For the observation at which concavity is imposed, the concavity test statistic $\hat{d}$ is 212 for the BC, 23 for the NQ, and 279 for the TL. It can be seen that the restricted and unrestricted results are increasingly different with the importance of $\hat{d}$. For the NQ specification, there are only relatively 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) found relatively small differences between the unrestricted and restricted estimates, whereas Gagné and Ouellette (1998) showed that the imposition of concavity can lead to important disparities between unrestricted and restricted estimates. 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 subsamples [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 that obtained on the pooled sample. There is some loss of precision in the BC estimates, however.

For the observation at which concavity is imposed, the concavity test statistic $\hat{d}$ is 76 for the BC, $\hat{d} = 5$ for the NQ and $\hat{d} = 261$ for the TL. In this light, it is not surprising that the disparity between the concavity-unrestricted and -restricted estimates is most important for the TL specification. In contrast 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 [i.e., 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 the demand for low-skilled labor. This contradicts the findings of most previous studies (see, e.g., the survey by Hamermesh 1993). Given the disparities that we have found, we must conclude that estimates of own-price elasticities are highly sensitive to the choice of functional form and sample split.

7.2 Cross-Price Elasticities

To measure factor substitution possibilities, we compute cross-price elasticities for the concavity-unrestricted and -restricted models. The tables for this section and the next, suppressed here, are available on request; see also the discussion paper version (Koebel et al. 2000). Here we discuss results only for the pooled sample and for sample split (a).

For the pooled sample BC unrestricted model, 18 out of 30 median cross-price elasticities are significant at the 5% level. This number reduces to 15 for the NQ and to 16 for the TL. The numbers of significant elasticities obtained from the split sample are 13, for the BC, 14 for the NQ, and 13 for the TL.

The cross-price elasticities computed on the basis of the NQ tend to be rather small in absolute value, pointing out a rigid production structure precluding frictionless substitution between inputs. In 22 out of 30 cases, the BC elasticities were greater than the corresponding NQ elasticities. For the TL, 21 elasticites are greater in absolute value than for the NQ, which seems to underestimates the extent of substitution and complementarity relationships.

Although there are some differences between restricted and unrestricted models, the number of contradictions is not very important. (By contradiction, we mean that an elasticity that is significantly different from 0 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. The choice of the functional form and sample split has then a greater impact on the estimates than the choice of whether or not to impose concavity.

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 can 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 United States and Canada (see Thompson and Taylor 1995). Because in our model, capital is assumed to be flexible, we adapt the definition given by Bergström and Panas (1992) and speak of capital–skill complementarity 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 skill. When $0 \leq \epsilon_{hp_k}$ it means that the degree of substitutability between labor and capital decreases with skill. 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. Our results differ somewhat from those of Falk and Koebel (2002) and Fitzenberger (1999) in the case where capital is quasi-fixed.

7.3 Output and Time Elasticities

Again we discuss output and time elasticities obtained from the pooled and split samples. These elasticities are in most cases significant at the 5% 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 (a) there are increasing returns to scale ($\epsilon_{cy} \leq 1$); (b) costs are reduced over time ($\epsilon_{ct} \leq 0$); (c) 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 1; (d) 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 that 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 level ($\epsilon_{hy} \geq \epsilon_{sy} \geq \epsilon_{uy}$). This result does not hold with the BC and the TL on the pooled sample. Sample split (a) provides 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 concavity-unrestricted and -restricted estimates.

7.4 Decomposition of Factor Demand Growth

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 toward 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 foregoing statistical tests that the BC model on sample split (a) is the specification providing the best overall fit. Because we are rather interested in the plausibility of the elasticities presented earlier, we follow an alternative approach in this 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. These effects can be identified from the total differentiation of the labor demand equations

$$\Delta g_{it}^* \simeq \sum_{j=k,h,s,u,e,m} \frac{\partial g^*}{\partial p_{jit}} \Delta p_{jit} + \frac{\partial g^*}{\partial y_{it}} \Delta y_{it} + \frac{\partial g^*}{\partial t}$$

$$\Leftrightarrow \frac{\Delta g_{it}^*}{g_{it}^*} \simeq \sum_{j=k,h,s,u,e,m} \varepsilon_{gp_j} \frac{\Delta p_{jit}}{p_{jit}} + \varepsilon_{gy} \frac{\Delta y_{it}}{y_{it}} + \varepsilon_{gt}, \quad (21)$$

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 easily calculated for each industry and time period. The different elasticities involved in (21) are computed for each industry and time period from the estimates. Then the predicted and observed values ($\Delta g_{it}^*/g_{it}^*$ and $\Delta g_{it}/g_{it}$) for each industry and time period are compared.

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 foregoing decomposition is based on a *first-order* approximation and is precise only for small changes Δp_{jit} and Δy_{it}. Although a second-order approximation would be more precise, *separate identification* of the impact of price, output, and time would then no longer be possible, because the second-order terms involve interacting variables.

Columns 3 and 5 of Table 6 give the observed and predicted changes 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% which is close to the prediction of 3.4% with the BC. From the comparisons of predicted and observed values across functional forms, we can conclude that the BC and the TL seem to be equally reliable models. The NQ appears to be the worst functional form; 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 foregoing criteria are inconclusive. Similarly, no firm conclusion emerges from the comparison between estimates on sample split (a) and on the pooled sample.

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 toward skilled labor and away from unskilled labor. The impact of the evolution of prices is (almost) always negative, but is also very small in absolute value (especially for the NQ). High-skilled labor is the labor input 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 in explaining the shift away from unskilled labor. For instance, only 0%–10% 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 Class

Functional form	Input demand	Actual change	Modeling assumptions	Predicted change	% Change attributable to		
					Price	Output	Time
Box–Cox	*h*	3.19	(a) unrestricted	3.37	−.64	1.31	2.70
			(a) restricted	3.32	−.68	1.18	2.82
			(p) unrestricted	3.80	−1.01	2.25	2.55
			(p) restricted	3.56	−1.04	2.13	2.48
	s	.30	(a) unrestricted	.47	−.24	1.21	−.50
			(a) restricted	.42	−.16	1.08	−.50
			(p) unrestricted	.35	−.25	1.23	−.62
			(p) restricted	.51	−.08	1.23	−.65
	u	−3.19	(a) unrestricted	−3.52	−.26	1.22	−4.48
			(a) restricted	−3.68	−.25	1.18	−4.61
			(p) unrestricted	−3.20	.02	1.56	−4.79
			(p) restricted	−3.21	−.01	1.55	−4.76
Normalized quadratic	*h*	3.19	(a) unrestricted	3.59	−.24	2.07	1.77
			(a) restricted	3.60	−.23	2.07	1.77
			(p) unrestricted	3.85	−.16	2.86	1.15
			(p) restricted	4.23	−.30	3.01	1.53
	s	.30	(a) unrestricted	.56	−.01	1.06	−.49
			(a) restricted	.60	−.01	1.06	−.45
			(p) unrestricted	.70	.04	1.03	−.38
			(p) restricted	.82	−.01	1.06	−.23
	u	−3.19	(a) unrestricted	−2.11	.08	.51	−2.70
			(a) restricted	−2.14	.05	.51	−2.70
			(p) unrestricted	−1.97	.03	.63	−2.62
			(p) restricted	−1.92	−.01	.64	−2.55
Translog	*h*	3.19	(a) unrestricted	2.81	−1.44	1.16	3.08
			(a) restricted	2.94	−1.26	1.16	3.03
			(p) unrestricted	3.13	−1.34	1.36	3.10
			(p) restricted	2.99	−1.46	1.14	3.31
	s	.30	(a) unrestricted	.65	−.37	1.40	−.39
			(a) restricted	.74	−.29	1.40	−.37
			(p) unrestricted	.32	−.37	1.20	−.51
			(p) restricted	.51	−.24	1.21	−.46
	u	−3.19	(a) unrestricted	−3.10	−.25	1.52	−4.37
			(a) restricted	−3.10	−.26	1.52	−4.36
			(p) unrestricted	−3.51	−.08	1.68	−5.11
			(p) restricted	−3.42	−.03	1.68	−5.07

NOTE: Column 3 shows the median growth rate over all industries and years. Columns 6–8 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 columns 6 to 8. The letter (a) denotes sample split (a), (p) denotes the pooled sample, "unrestricted" represents the concavity unrestricted specification and "restricted" represents for the concavity restricted specification.

8. CONCLUSION

In this article we have proposed a method for imposing curvature conditions on a wide class of functional forms and for testing these restrictions. In the empirical application, we estimated the parameters of a concavity-constrained BC cost function. For our dataset, a parametric test for the null of concavity leads to a weak rejection of this assumption.

Because 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 relationship 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, the impact of output and time is in general more important than that of price and substitution. 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 cannot, however, explain much of the observed changes in the different types of labor inputs. Whereas the evolution of skilled labor demand is explained mainly by output growth, the dominant factor "explaining" the shift against unskilled labor and toward high-skilled labor is the residual time trend. This emphasizes the necessity to extend the usual theoretical framework in production analysis to promote a better understanding of technologic change and its determinants.

ACKNOWLEDGMENTS

We greatly benefited from the constructive recommendations of an associate editor and two anonymous referees. Thanks also to Herbert Buscher, Sandra Kneile, and participants in seminars in Lausanne, London, Magdeburg, Mannheim, Strasbourg, Tilburg, and Tübingen for useful comments. Any shortcomings and errors are, of course, our own responsibility.

APPENDIX A: PROOFS OF THE PROPOSITIONS

Proof of Proposition 1

For the sake of completeness, we adapt the results of Gouriéroux et al. (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.

Parts a and b. It is clear that $\hat{\alpha}$ and $\tilde{\alpha}^0$ converges to the true parameter α_0 under H_0. It remains to show that $\hat{\alpha}^0$ also converges to α_0. For this purpose, we characterize in stages 1–3 the necessary conditions for an optimum in (9), (10), and (12), and show in stage 4 that the conditions corresponding to (9) are equivalent to those corresponding to (10) and (12).

Stage 1. Problem (9) can be reparameterized using the Cholesky decomposition to transform the inequality constraints $v'\underline{H}v \leq 0$, $\forall v$, into the *equality constraints* $g(\alpha) = \eta_H^0(u)$. The corresponding Lagrangean is

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

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

$$\frac{\partial\mathcal{L}}{\partial\alpha} = 0 \Leftrightarrow -2\widehat{\Omega}_\alpha^{-1}(\hat{\alpha}-\tilde{\alpha}^0) - \frac{\partial g'}{\partial\alpha}(\tilde{\alpha}^0)\tilde{\mu}^0 = 0, \qquad \text{(A.1)}$$

$$\frac{\partial\mathcal{L}}{\partial u} = 0 \Leftrightarrow \frac{\partial\eta_H^{0\prime}}{\partial u}(\tilde{u}^0)\tilde{\mu}^0 = 0, \qquad \text{(A.2)}$$

and

$$\frac{\partial\mathcal{L}}{\partial\mu} = 0 \Leftrightarrow \eta_H^0(\tilde{u}^0) = g(\tilde{\alpha}^0). \qquad \text{(A.3)}$$

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

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

and

$$(\hat{\alpha}-\tilde{\alpha}^0)'\widehat{\Omega}_\alpha^{-1}(\hat{\alpha}-\tilde{\alpha}^0) = \frac{\tilde{\mu}^{0\prime}}{2}\frac{\partial g}{\partial\alpha'}(\tilde{\alpha}^0)\widehat{\Omega}_\alpha\frac{\partial g'}{\partial\alpha}(\tilde{\alpha}^0)\frac{\tilde{\mu}^0}{2}. \qquad \text{(A.5)}$$

Stage 2. The Lagrangean corresponding to (12) is given by

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

and the solution $(\hat{\alpha}^0, \hat{\mu}^0)$ satisfies

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

and

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

Similarly to the former problem, we obtain

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

and

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

Stage 3. The first-order conditions for a solution to (10) lead to

$$\frac{\partial\eta_H^{0\prime}}{\partial u}(\hat{u}^0)\widehat{\Omega}_H^{-1}\big(\hat{\eta}_H - \eta_H^0(\hat{u}^0)\big) = 0, \qquad \text{(A.10)}$$

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

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

which is asymptotically equivalent to [using (A.8) and (11)]

$$\frac{\partial\eta_H^{0\prime}}{\partial u}(\hat{u}^0)\hat{\mu}^0 \stackrel{a}{=} 0. \qquad \text{(A.11)}$$

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

$$-2\widehat{\Omega}_\alpha^{-1}(\hat{\alpha}-\hat{\alpha}^0) - \frac{\partial g'}{\partial\alpha}(\hat{\alpha}^0)\hat{\mu}^0 = 0,$$

$$\eta_H^0(\hat{u}^0) - g(\hat{\alpha}^0) = 0,$$

and

$$\frac{\partial\eta_H^{0\prime}}{\partial u}(\hat{u}^0)\hat{\mu}^0 \stackrel{a}{=} 0.$$

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

Part c. As $\hat{\alpha}$, $\tilde{\alpha}^0$, and $\hat{\alpha}^0$ converge to the true parameter α_0 under H_0 (part a), we directly see from (A.4)–(A.5) and (A.8)–(A.9) of part a, that the minima of (9) and (12) are asymptotically equivalent. We now show that $(\hat{\alpha} - \hat{\alpha}^0)'\widehat{\Omega}_\alpha^{-1}(\hat{\alpha}-\hat{\alpha}^0)$ is asymptotically equivalent to $(\hat{\eta}_H - \hat{\eta}_H^0)'\widehat{\Omega}_H^{-1}(\hat{\eta}_H - \hat{\eta}_H^0)$ under H_0, with $\hat{\eta}_H^0 \equiv \eta_H^0(\hat{u}^0) = g(\hat{\alpha}^0)$. Using a first-order Taylor development of $g(\hat{\alpha}^0)$ around $\hat{\alpha}$, we can write, under H_0,

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

where the second equality follows from (A.6) and the last equality follows from (A.9).

Part d. Using a first-order Taylor expansion of (A.6)–(A.7) around $\hat{\alpha}$, we can rewrite these conditions as

$$\hat{\alpha}^0 - \hat{\alpha} \overset{a}{=} \widehat{\Omega}_\alpha \frac{\partial g}{\partial \alpha}(\hat{\alpha})' \frac{\hat{\mu}^0}{2}$$

and

$$\hat{\eta}_H^0 - g(\hat{\alpha}) - \frac{\partial g}{\partial \alpha'}(\hat{\alpha})(\hat{\alpha}^0 - \hat{\alpha}) \overset{a}{=} 0.$$

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

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

and, finally,

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

Proof of Proposition 2

To ensure that the minimization in (14) occurs on the domain where $\lambda \le 0$, 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

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

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

$$Q'\overline{H}Q = \Lambda.$$

From this equation, it follows that (see, e.g., Magnus 1985; Kodde and Palm 1987)

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

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

$$\begin{aligned} \widehat{\Sigma}^- &= \left[\frac{\partial \lambda}{\partial \text{vec}'(\overline{H})}(\widehat{\overline{H}})\widehat{\mathbf{\Omega}}_H \frac{\partial \lambda'}{\partial \text{vec}(\overline{H})}(\widehat{\overline{H}})\right]^- \\ &= \left[P(\widehat{Q}' \otimes \widehat{Q}')\widehat{\mathbf{\Omega}}_H(\widehat{Q}' \otimes \widehat{Q}')'P'\right]^- \\ &= P(\widehat{Q}' \otimes \widehat{Q}')\widehat{\mathbf{\Omega}}_H^-(\widehat{Q}' \otimes \widehat{Q}')'P', \quad \text{(A.13)} \end{aligned}$$

using the orthogonality of $(\widehat{Q}' \otimes \widehat{Q}')$ and $(\widehat{Q}' \otimes \widehat{Q}')'$ and the fact that PP' is an identity matrix. It is then easy to check that, denoting $\widehat{\Sigma} = P(\widehat{Q}' \otimes \widehat{Q}')\widehat{\mathbf{\Omega}}_H(\widehat{Q}' \otimes \widehat{Q}')'P'$, one indeed has $\widehat{\Sigma}\widehat{\Sigma}^-\widehat{\Sigma} = \widehat{\Sigma}$. Note that $\widehat{\Omega}_H$, the variance matrix of $\hat{\eta}_H \equiv \text{vecli}\widehat{\overline{H}}$, is a submatrix of $\widehat{\mathbf{\Omega}}_H$. Let $\overline{H}^0 = \widehat{Q}\Lambda^0\widehat{Q}'$; using (A.12) and (A.13), we can rewrite

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

where the third equality follows from the properties of the vec operator and Kronecker product, and the last from Dhrymes's (1994) lemma A1. Because $\overline{H}^0 = \widehat{Q}\Lambda^0\widehat{Q}'$ is negative semidefinite, it can be written as $-U'U$. However, $\widehat{Q}\Lambda^0\widehat{Q}'$ comprises only the $S_p - 1$ free parameters of ν, whereas $-U'U$ is composed of the $S_p(S_p - 1)/2$ free parameters u. Hence, in small samples,

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

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

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

Proof of Proposition 3

a. An eigenvalue λ_1 of $\overline{H}$ is a solution of $f(\lambda, \overline{H}) \equiv |\lambda I_{S_p} - \overline{H}| = 0$. 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 $\partial f(\lambda, \overline{H})/\partial\lambda \ne 0$ at $\lambda = \lambda_1$. Because

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

it follows that

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

Hence

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

Thus, $\partial f(\lambda, \overline{H})/\partial\lambda$ is different from 0 if and only if λ_1 is simple. A related result has been obtained by Magnus (1985, thm. 1).

b. We adapt a proof given by Lau (1978, lems. 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 0. Because 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 0.

APPENDIX B: SOME FUNCTIONAL FORMS NESTED WITHIN THE BOX–COX SPECIFICATION

Here we show how the NQ, a version of the GL, and the TL are obtained as special cases of the generalized BC specification. The derivation of further interesting functional forms can also be obtained along these lines.

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

$$Z_{it} = z_{it} - \iota_{S_z}$$

and

$$P_{it} = \frac{p_{it}}{p'_{it}\theta_i} - \iota_{S_p},$$

where ι_{S_p} is a S_p-vector of 1s. The cost function (15) then becomes

$$\begin{aligned}
c^* &= p'_{it}\bar{x}_i C^*_{NQ}(p_{it}, z_{it}; \beta_i) + (p'_{it}\bar{x}_i)\\
&= p'_{it}\bar{x}_i\Big(1 + \beta_{0i} + P'_{it}B_{pi} + Z'_{it}B_z + \frac{1}{2}P'_{it}B_{pp}P_{it} + P'_{it}B_{pz}Z_{it}\\
&\qquad + \frac{1}{2}Z'_{it}B_{zz}Z_{it}\Big)\\
&= p'_{it}\bar{c}_iB_{pi} + (p'_{it}\bar{x}_i)z'_{it}B_z + \frac{1}{2}\bar{c}_i\frac{p'_{it}B_{pp}p_{it}}{p'_{it}\theta_i} + \bar{c}_ip'_{it}B_{pz}z_{it}\\
&\quad + (p'_{it}\bar{x}_i)\frac{1}{2}z'_{it}B_{zz}z_{it} + p'_{it}\bar{x}_i\Big(1 + \beta_{0i} - \iota'_{S_p}B_{pi} - \iota'_{S_z}B_z\\
&\qquad + \frac{1}{2}\iota'_{S_p}B_{pp}\iota_{S_p} + \iota'_{S_p}B_{pz}\iota_{S_z} + \frac{1}{2}\iota'_{S_z}B_{zz}\iota_{S_z}\Big)\\
&\quad + p'_{it}\bar{x}_i\Big(-\iota'_{S_p}B_{pp}\frac{p_{it}}{p'_{it}\theta_i} - \iota'_{S_p}B_{pz}z_{it} - \frac{p'_{it}}{p'_{it}\theta_i}B_{pz}\iota_{S_z} - \iota'_{S_z}B_{zz}z_{it}\Big).
\end{aligned}$$

Considering the restrictions (17), we obtain

$$\begin{aligned}
c^* &= p'_{it}\bar{c}_iB_{pi} + (p'_{it}\bar{x}_i)z'_{it}B_z + \frac{1}{2}\bar{c}_i\frac{p'_{it}B_{pp}p_{it}}{p'_{it}\theta_i} + \bar{c}_ip'_{it}B_{pz}z_{it}\\
&\quad + \frac{1}{2}(p'_{it}\bar{x}_i)z'_{it}B_{zz}z_{it} + p'_{it}\bar{x}_i\Big(\beta_{0i} - \iota'_{S_z}B_z + \frac{1}{2}\iota'_{S_z}B_{zz}\iota_{S_z}\Big)\\
&\quad + p'_{it}\bar{x}_i\Big(-\frac{p'_{it}}{p'_{it}\theta_i}B_{pz}\iota_{S_z} - \iota'_{S_z}B_{zz}z_{it}\Big)\\
&= p'_{it}\mathbf{B}_{pi} + (p'_{it}\bar{x}_i)z'_{it}\mathbf{B}_z + \frac{1}{2}\frac{p'_{it}\mathbf{B}_{ppi}p_{it}}{p'_{it}\theta_i} + p'_{it}\mathbf{B}_{pzi}z_{it}\\
&\quad + (p'_{it}\bar{x}_i)\frac{1}{2}z'_{it}B_{zz}z_{it},
\end{aligned}$$

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

$$\mathbf{B}_{pi} = \bar{c}_i(B_{pi} - B_{pz}\iota_{S_z}) + \bar{x}_i\Big(\beta_{0i} - \iota'_{S_z}B_z + \frac{1}{2}\iota'_{S_z}B_{zz}\iota_{S_z}\Big),$$

$$\mathbf{B}_z = B_z - B_{zz}\iota_{S_z},$$

$$\mathbf{B}_{ppi} = B_{pp}\bar{c}_i,$$

and

$$\mathbf{B}_{pzi} = B_{pz}\bar{c}_i.$$

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

$$Z_{it} = 2\big(z_{it}^{1/2} - \iota_{S_z}\big)$$

and

$$P_{it} = 2\left[\left(\frac{p_{it}}{p'_{it}\theta_i}\right)^{1/2} - \iota_{S_p}\right],$$

where, by convention, $z_{it}^{1/2} = (z_1^{1/2}, \ldots, z_{S_z}^{1/2})'_{it}$ and $p_{it}^{1/2} = (p_1^{1/2}, \ldots, p_{S_p}^{1/2})'_{it}$. The cost function (15) then becomes

$$\begin{aligned}
c^* &= p'_{it}\bar{x}_iC^*_{GL}(p_{it}, z_{it}; \beta_i) + p'_{it}\bar{x}_i\\
&= p'_{it}\bar{x}_i\Big(1 + \beta_{0i} + P'_{it}B_{pi} + Z'_{it}B_z + \frac{1}{2}P'_{it}B_{pp}P_{it} + P'_{it}B_{pz}Z_{it}\\
&\qquad + \frac{1}{2}Z'_{it}B_{zz}Z_{it}\Big)\\
&= 2\bar{c}_i(p'_{it}\theta_i)^{1/2}p_{it}^{1/2'}B_{pi} + 2(p'_{it}\bar{x}_i)z_{it}^{1/2'}B_{zi} + 2\bar{c}_ip_{it}^{1/2'}B_{pp}p_{it}^{1/2}\\
&\quad + 4\bar{c}_i(p'_{it}\theta_i)^{1/2}p_{it}^{1/2'}B_{pz}z_{it}^{1/2} + 2(p'_{it}\bar{x}_i)z_{it}^{1/2'}B_{zz}z_{it}^{1/2}\\
&\quad + 2p'_{it}\bar{x}_i\big(1/2 + \beta_{0i} - \iota'_{S_p}B_{pi} - \iota'_{S_z}B_z + \iota'_{S_p}B_{pp}\iota_{S_p}\\
&\qquad + 2\iota'_{S_p}B_{pz}\iota_{S_z} + \iota'_{S_z}B_{zz}\iota_{S_z}\big)\\
&\quad + 4p'_{it}\bar{x}_i\Bigg[-\iota'_{S_p}B_{pp}\left(\frac{p_{it}}{p'_{it}\theta_i}\right)^{1/2} - \iota'_{S_p}B_{pz}z_{it}^{1/2}\\
&\qquad - \left(\frac{p'_{it}}{p'_{it}\theta_i}\right)^{1/2}B_{pz}\iota_{S_z} - \iota'_{S_z}B_{zz}z_{it}^{1/2}\Bigg]\\
&= 2\bar{c}_i(p'_{it}\theta_i)^{1/2}p_{it}^{1/2'}B_{pi} + 2(p'_{it}\bar{x}_i)z_{it}^{1/2'}B_z + 2\bar{c}_ip_{it}^{1/2'}B_{pp}p_{it}^{1/2}\\
&\quad + 4\bar{c}_i(p'_{it}\theta_i)^{1/2}p_{it}^{1/2'}B_{pz}z_{it}^{1/2} + 2(p'_{it}\bar{x}_i)z_{it}^{1/2'}B_{zz}z_{it}^{1/2}\\
&\quad + 2p'_{it}\bar{x}_i(\beta_{0i} - 1/2 - \iota'_{S_z}B_z + \iota'_{S_z}B_{zz}\iota_{S_z})\\
&\quad + 4p'_{it}\bar{x}_i\left[-\left(\frac{p'_{it}}{p'_{it}\theta_i}\right)^{1/2}B_{pz}\iota_{S_z} - \iota'_{S_z}B_{zz}z_{it}^{1/2}\right].
\end{aligned}$$

After reparameterization we obtain

$$\begin{aligned}
c^* &= (p'_{it}\theta_i)^{1/2}p_{it}^{1/2'}\mathbf{B}_{pi} + (p'_{it}\bar{x}_i)z_{it}^{1/2'}\mathbf{B}_z + p_{it}^{1/2'}\mathbf{B}_{ppi}p_{it}^{1/2}\\
&\quad + (p'_{it}\theta_i)^{1/2}p_{it}^{1/2'}\mathbf{B}_{pzi}z_{it}^{1/2} + (p'_{it}\bar{x}_i)z_{it}^{1/2'}\mathbf{B}_{zz}z_{it}^{1/2},
\end{aligned}$$

which is a version of the GL cost function, with

$$\mathbf{B}_{pi} = \bar{c}_i(2B_{pi} - 4B_{pz}\iota_{S_z}),$$
$$\mathbf{B}_z = 2B_z - 4B_{zz}\iota_{S_z},$$
$$\mathbf{B}_{ppi} = 2B_{pp}\bar{c}_i + \mathrm{diag}\left[2\bar{x}_i(\beta_{0i} - 1/2 - \iota'_{S_z}B_z + \iota'_{S_z}B_{zz}\iota_{S_z})\right],$$
$$\mathbf{B}_{pzi} = 4B_{pz}\bar{c}_i,$$

and

$$\mathbf{B}_{zz} = 2B_{zz},$$

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

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

$$Z_{it} = \lim_{\gamma_1 \to 0} \frac{z_{it}^{\gamma_1} - 1}{\gamma_1} = \ln z_{it},$$
$$C^*_{TL} = \lim_{\gamma_2 \to 0} \frac{(c^*/p'_{it}\bar{x}_i)^{\gamma_2} - 1}{\gamma_2} = \ln c^* - \ln(p'_{it}\bar{x}_i),$$
$$P_{it} = \lim_{\gamma_1 \to 0} \frac{(p_{it}/p'_{it}\theta_i)^{\gamma_1} - \iota_{S_p}}{\gamma_1} = \ln p_{it} - \iota_{S_p}\ln(p'_{it}\theta_i),$$

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

$$\begin{aligned}
\ln c^* &= C^*_{TL}(p_{it}, z_{it}; \beta_i) + \ln(p'_{it}\bar{x}_i) \\
* &= \Big(\beta_{0i} + P'_{it}B_{pi} + Z'_{it}B_{zi} + \frac{1}{2}P'_{it}B_{pp}P_{it} + P'_{it}B_{pz}Z_{it} \\
&\quad + \frac{1}{2}Z'_{it}B_{zz}Z_{it}\Big) + \ln(p'_{it}\bar{x}_i) \\
&= \beta_{0i} + (\ln p_{it})'B_{pi} + (\ln z_{it})'B_{zi} + \frac{1}{2}(\ln p_{it})'B_{pp}(\ln p_{it}) \\
&\quad + (\ln p_{it})'B_{pz}\ln z_{it} + \frac{1}{2}(\ln z_{it})'B_{zz}\ln z_{it} + \ln(p'_{it}\bar{x}_i) \\
&\quad + \ln(p'_{it}\theta_i)\Big(-\iota'_{S_p}B_{pi} + \tfrac{1}{2}\ln(p'_{it}\theta_i)\iota'_{S_p}B_{pp}\iota_{S_p} \\
&\quad - \iota'_{S_p}B_{pp}(\ln p_{it}) - \iota'_{S_p}B_{pz}\ln z_{it}\Big). \qquad \text{(B.1)}
\end{aligned}$$

Once restrictions (17) are imposed, the last line of (B.1) boils down to $\ln(\bar{c}_i)$ and the usual TL specification is obtained by subsuming $\ln(\bar{c}_i)$ into β_{0i}.

APPENDIX C: DESCRIPTION OF THE INDUSTRIES

Table C.1. Denomination of the Industries

No.	Industry	No.	Industry
14	Chemical products	30	Aircraft and spacecraft
15	Refined petroleum products	31	Electrical machinery, equipment and appliances
16	Plastic products	32	Precision instruments and optical equipment
17	Rubber products	33	Tools and finished metal products
18	Quarrying, building materials, etc.	34	Musical instruments, games and toys, jewelry, etc.
19	Ceramic products	35	Wood working
20	Glass products	36	Wood products
21	Iron and steel	37	Pulp, paper, and paperboard
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 (excluding beverages)
28	Road vehicles	44	Beverages
29	Ships and boats	45	Tobacco products

Table C.2. Description of the Different Subsamples Considered

	Subsample (i)	Subsample (ii)	Subsample (iii)
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 industries	Medium industries	Large industries
Industry no.	17, 19, 20, 23, 29, 30, 34, 35, 37, 38, 40	22, 24, 25, 27, 32, 36, 39, 42, 44, 45	14, 16, 18, 21, 26, 28, 31, 33, 41, 43
Split (c)	Low skill intensive	Skill intensive	High skill intensive
Industry no.	22, 24, 34, 35, 36, 38, 40, 41, 42, 43, 44	16, 18, 20, 21, 23, 28, 33, 37, 39, 45	14, 17, 19, 25, 26, 27, 29, 30, 31, 32
Split (d)	Not labor intensive	Labor intensive	Highly labor intensive
Industry no.	14, 18, 21, 22, 35, 37, 38, 40, 43, 44, 45	16, 17, 20, 24, 27, 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, 30, 31, 32, 33, 36, 40, 42, 43	14, 17, 24, 28, 29, 33, 34, 38, 41, 45	18, 19, 20, 21, 23, 27, 35, 37, 39, 44

[Received May 1999. Revised January 2002.]

REFERENCES

Baltagi, B. H. (1995), *Econometric Analysis of Panel Data*, New York: Wiley.

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

Bergström, V., and Panas, E. E. (1992), "How Robust Is the Capital-Skill Complementarity Hypothesis?," *Review of Economics and Statistics,* 74, 540–552.

Berman, E., Bound, J., and Griliches, Z. (1994), "Changes in the Demand for Skilled Labor within U.S. Manufacturing, Evidence From the Annual Survey of Manufactures," *Quarterly Journal of Economics,* 109, 367–397.

Berndt, E. R., and Khaled, M. S. (1979), "Parametric Productivity Measurement and Choice Among Flexible Functional Forms," *Journal of Political Economy,* 87, 1220–1245.

Betts, J. R. (1997), "The Skill Bias of Technological Change in Canadian Manufacturing Industries," *Review of Economics and Statistics,* 79, 146–150.

Dhrymes, P. J. (1994), "Specification Tests in Simultaneous Equation Systems," *Journal of Econometrics*, 64, 45–76.

Diewert, W. E., and Wales, T. J. (1987), "Flexible Functional Forms and Global Curvature Conditions," *Econometrica,* 55, 47–68.

Falk, M., and Koebel, B. M. (2002), "Outsourcing, Imports and Labour Demand," *Scandinavian Journal of Economics,* 104, 567–586.

Fitzenberger, B. (1999), *Wages and Employment Across Skill Groups: An Analysis for West Germany*, ZEW Economic Studies 6, Heidelberg: Physica Verlag.

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

Gallant, A. R., and Golub, G. H. (1984), "Imposing Curvature Restrictions on Flexible Functional Forms," *Journal of Econometrics*, 26, 295–321.

Gouriéroux, C., Holly, A., and Monfort, A. (1982), "Likelihood Ratio Test, Wald, and Kuhn–Tucker Test in Linear Models With Inequality Constraints on the Regression Parameters," *Econometrica*, 50, 63–80.

Gouriéroux, C., and Monfort, A. (1989), *Statistique et Modèles Économétriques*, Paris: Economica.

Gouriéroux, C., Monfort, A., and Trognon, A. (1985), "Moindres Carrés Asymptotiques," *Annales de l'INSEE*, 58, 91–122.

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

Härdle, W., and Hart, J. D. (1992), "A Bootstrap Test for Positive Definiteness of Income Effect Matrices," *Econometric Theory*, 8, 276–290.

Härdle, W., Hildenbrand, W., and Jerison, M. (1991),"Empirical Evidence on the Law of Demand," *Econometrica*, 59, 1525–1549.

Kodde, D. A., and Palm, F. C. (1986), "Wald Criteria for Jointly Testing Equality and Inequality Restrictions," *Econometrica*, 54, 1243–1248.

——— (1987), "A Parametric Test of the Negativity of the Substitution Matrix," *Journal of Applied Econometrics,* 2, 227–235.

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

Koebel, B. M. (1998), "Tests of Representative Firm Models: Results for German Manufacturing Industries," *Journal of Productivity Analysis*, 10, 251–270.

——— (2002), "Can Aggregation be Achieved by Neglecting the Problem? Property Inheritance and Aggregation Biases," *International Economic Review*, 43, 223–255.

Koebel, B. M., Falk, M., and Laisney, F. (2000), "Imposing and Testing Curvature Conditions on a Box–Cox Cost Function," Discussion Paper 00-70, ZEW (download at *http://www.zew.de*).

Lansink, A. O., and Thijssen, G. (1998), "Testing Among Functional Forms: An Extension of the Generalized Box–Cox Formulation," *Applied Economics*, 30, 1001–1010.

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

Magnus, J. R. (1985), "On Differentiating Eigenvalues and Eigenvectors," *Econometric Theory*, 1, 179–191.

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

Ruud, P. A. (1997), "Restricted Least Squares Subject to Monotonicity and Concavity Constraints," in *Advances in Economic and Econometrics: Theory and Applications*, Vol. III, eds. D. M. Kreps and K. F. Wallis, New York: Cambridge University Press.

Ryan, D. L., and Wales, T. J. (1998), "A Simple Method for Imposing Local Curvature in Some Flexible Consumer-Demand Systems," *Journal of Business & Economic Statistics*, 16, 331–337.

——— (2000), "Imposing Local Concavity in the Translog and Generalized Leontief Cost Functions," *Economic Letters*, 67, 253–260.

Terrell, D. M. (1996), "Incorporating Monotonicity and Concavity Conditions in Flexible Functional Forms," *Journal of Applied Econometrics*, 11, 179–194.

Thompson, P., and Taylor, T. G. (1995), "The Capital-Energy Substitutability Debate: A New Look," *Review of Economics and Statistics,* 77, 565–569.

Wolak, F. A. (1989), "Local and Global Testing of Linear and Nonlinear Inequality Constraints in Nonlinear Econometric Models," *Econometric Theory*, 5, 1–35.

Wooldridge, J. M. (1992), "Some Alternatives to the Box–Cox Regression Model," *International Economic Review*, 33, 935–955.