ETLA ELINKEINOELÄMÄN TUTKIMUSLAITOS
THE RESEARCH INSTITUTE OF THE FINNISH ECONOMY
Lönnrotinkatu 4 B, 00120 Helsinki 12, Finland, tel. 601322

Keskusteluaiheita
Discussion papers

Pekka Ilmakunnas

STOCHASTIC CONSTRAINTS ON COST FUNCTION PARAMETERS: MIXED AND HIERARCHICAL APPROACHES*

N:o 181 31.10.1985

*Forthcoming in Empirical Economics.

Abstract

Symmetry and homogeneity constraints are imposed stochastically in a system of cost share equations. This is done using both mixed information and hierarchical representations. The approaches are compared in a Monte Carlo study and using data from U.S. manufacturing.

I Introduction

Economic theory often implies constraints on the parameters of a set of behavioral equations, both within and across equations. Obvious examples are homogeneity and symmetry constraints in demand systems. To obtain estimates consistent with the behavioral assumption, these constraints should be imposed in estimation. Alternatively, one may want to test the constraints by estimating the model in both constrained and unconstrained form. There are, however, some reasons for allowing the constraints not to hold exactly. First of all, the type of data usually available e.g. in demand or cost studies is rather aggregate and hence the same constraints that hold for micro units may no longer hold. Hence, testing the constraints or imposing them exactly may be inappropriate.[2)]

Second, the optimizing behavior that leads to the parameter constraints may be imperfect. A common justification for adding an error term in behavioral equations is optimization errors, which lead to first-order conditions for the optimum not to hold exactly. By the same token, one may want to treat the parameter constraints stochastic due to optimization errors. Third, varying the a priori strictness of the constraints allows us to study how fragile the results are to imposition of the constraints and hence, how useful the point estimates are.[3)] Finally, we may note that some recent work has shown that the standard tests may be biased towards rejecting hypotheses on parameter constraints in large demand systems.[4)]

Stochastic prior information of parameter values and constraints has been taken into account in the work on consumer demand systems where occasionally Theil-Goldberger (1961) mixed estimation has been used (e.g. Paulus (1975)). Kiefer (1977), on the other hand, used the hierarchical model to impose constraints stochastically in a demand system. Kiefer's approach has been used in a cost function context by Ilmakunnas (1985) who also estimated underlying second-stage parameters and in Tsurumi, Wago and Ilmakunnas (1985). For the first-stage parameters the constraints hold stochastically, whereas in the second stage they are exact.

The purpose to this paper is to compare the mixed and hierarchical approaches for specifying parameter constraints, using estimation of a cost share system as an example. We will study how the demand and substitution elasticities change when the a priori strictness of the constraints is changed. In a Monte Carlo study we compare the approaches in terms of the mean squared errors of the estimated parameters.

II The model and the estimation methods

In this section we will briefly describe the model studied and the alternative ways of imposing the constraints. We consider the estimation of a KLEM translog unit cost function

$$\ln C = \alpha_0 + \sum_i \alpha_i \ln w_i + \frac{1}{2}\sum_i\sum_j \beta_{ij} \ln w_i \ln w_j \tag{1}$$

where C is average total cost and w_i (i=K,L,E,M) is the price of input i. By Shephard's lemma we obtain a system of cost share equations

$$S_i = \alpha_i + \sum_j \beta_{ij} \ln w_j \qquad i=K,L,E,M. \tag{2}$$

The following linear parameter constraints are implied by the properties of cost functions:

$$\beta_{ij} = \beta_{ji} \qquad \text{(symmetry)} \tag{3}$$

$$\sum_i \alpha_i = 1, \quad \sum_j \beta_{ij} = 0 \qquad \text{(homogeneity)} \tag{4}$$

$$\sum_i \alpha_i = 1, \quad \sum_i \beta_{ij} = 0 \qquad \text{(adding-up)} \tag{5}$$

In addition, inequality constraints are implied by the concavity and monotonicity of cost functions. We will, however, not consider the latter constraints in this paper since they cannot be expressed as linear equalities. The traditional stochastic specification of the model is to add an error term to (2) to reflect optimization errors.

Due to the adding-up constraint the system is singular; hence the materials share equation is dropped. The vector of errors of the remaining three equations is assumed to be normally distributed with zero vector mean and covariance matrix $\Sigma \otimes I$.

The constraints (3) and (4) can be imposed stochastically in two different ways. Define the T x 5 data matrix X with t -th row $(1, \ln w_{Kt}, \ln w_{Lt}, \ln w_{Et}, \ln w_{Mt})$ and a 1 x 3T vector of cost shares $y' = (S_{K1}, \ldots, S_{ET})$. Finally define a 1 x 15 vector of parameters $\beta' = (\alpha_K, \beta_{KK}, \beta_{KL}, \beta_{KE}, \beta_{KM}, \alpha_L, \ldots, \beta_{EM})$.

The unconstrained cost share system can be written as

$$y = (I \otimes X)\beta + e \quad , \quad e \sim N(0, \Sigma \otimes I) \quad . \tag{6}$$

In the mixed approach we specify the constraints in equation

$$R\beta = u \quad , \quad u \sim N(0, \Phi) \tag{7}$$

where R is a 6 x 15 matrix. The constraints have the form $\beta_{KL} - \beta_{LK} = u_1$, and hence $E(\beta_{KL} - \beta_{LK}) = 0$, etc. The mixed estimator of β is, assuming Φ and σ_R^2 known,

$$b_R = (\Sigma^{-1} \otimes X'X + R'\Phi^{-1}R)^{-1}(\Sigma^{-1} \otimes X')y \ . \tag{8}$$

Mixed estimation has been criticized by Bayesians (e.g. Zellner (1975), Swamy and Mehta (1983)) on the grounds that R is considered fixed although u is stochastic in (7). One can, however, give the method a Bayesian interpretation (e.g. Theil (1971), pp. 670-2), using (7) as the prior distribution of the parameters.

In the hierarchical approach, (6) is treated as a first-stage model. Its parameters β depend on hyperparameters γ through the second-stage model, which can be interpreted as a prior distribution of β in a Bayesian approach:

$$\beta = Q\gamma + v \qquad , \qquad v \sim N(0,\Omega) \tag{9}$$

where Q is a 15 x 9 matrix such that RQ = 0 (see Evans and Patterson (1985)) and γ is a 9 x 1 vector. γ can depend on third-stage parameters, but here diffuse prior information on γ is assumed. The parameter constraints hold for the expectation of a linear combination of elements of γ. For example, $\beta_{KL} = \gamma_3 + v_3$ and $\beta_{LK} = \gamma_3 + v_7$; hence $E\beta_{KL} = E\beta_{LK} = E\gamma_3$. Similarly, $\beta_{KM} = -\gamma_2 - \gamma_3 - \gamma_4 + v_5$ and $E\beta_{KM} = E(-\gamma_2 - \gamma_3 - \gamma_4) = E(-\beta_{KK} - \beta_{KL} - \beta_{KE})$ etc. Because of the parameter restrictions, only nine second-stage parameters are needed.

Assuming that Σ and Ω are known it can be shown, (see Ilmakunnas (1985)) that β and γ are normally distributed with means respectively,

$$b_Q = \{\Sigma^{-1} \otimes X'X + \Omega^{-1} - \Omega^{-1}Q(Q'\Omega^{-1}Q)^{-1}Q'\Omega)\}^{-1}(\Sigma^{-1} \otimes X')y \tag{10}$$

and

$$g = \{Q'(\Sigma \otimes (X'X)^{-1} + \Omega)^{-1}Q\}^{-1}Q'\{\Sigma \otimes (X'X)^{-1} + \Omega\}^{-1}b \tag{11}$$

where $b = (I \otimes (X'X)^{-1}X')y$ is the OLS estimate of β. An alternative form for b_Q is, using the lemma in Smith (1973),

$$b_Q = (\Sigma^{-1} \otimes X'X + \Omega^{-1})^{-1} \{(\Sigma^{-1} \otimes X'X)b + \Omega^{-1}Qg\} \qquad (12)$$

which shows that b_g is a weighed average of the OLS estimate b and the second-stage estimates g.

The hierarchical model was originally suggested by Lindley and Smith (1972) and Smith (1973) in a Bayesian framework, where (9) is treated as the prior probability distribution of the parameters β. However, the method can also be given a sampling theoretic interpretation (Haitovsky (1979)).

The hierarchical model is also closely related to random coefficient models. They are often used in cross-section models or in pooling to control for heteroscedasticity. Swamy's (1970) random coefficient model corresponds in the Lindley-Smith approach to within-equation exchangeability of the parameters, i.e. the parameter vectors of different cross-sectional equations have the same mean vector. Lindley and Smith (1972) also discuss within-equation exchangeability, i.e. the parameters of an equation have the same mean. They show this to be similar to ridge regression estimation. The approach suggested by Kiefer (1977) and followed here is a variant of the model where exchangeability is introduced only through the constraints on the parameters.

Comparison of (8) and (10) shows that in both cases the estimate of β gives different values for β_{ij} and β_{ji}, reflecting the stochastic nature of the symmetry constraint. On the other hand, the hierarchical approach gives, in addition, second-stage parameters γ. As second-stage estimates of the cost function parameters we can therefore use the vector Qg, where constraints (3) and (4) hold.

The choice between the two approaches depends partly on what is assumed to be stochastic in the model and on how much prior information of the parameter values is available. In the mixed model, as presented in (7), the stochasticity is restricted to the constraints. The parameters themselves are not explicitly stochastic. In particular, the parameters that do not appear in the constraints, i.e. the α's, are non-stochastic. If they are made stochastic, we have to specify prior means by adding a vector r in the right hand side of (7) and modifying R accordingly. In contrast, in the hierarchical model (9) all parameters are explicitly stochastic, i.e. they have their own error terms. If we had prior information on the means of the parameters, we could replace some of the elements of γ by the a priori values. However, if no a priori values are available, we can estimate parameters γ using (11). In the mixed model this is not possible. In this sense the hierarchical model is more flexible in representing parameter stochasticity.

A problem with both approaches is the specification of the prior covariances of the parameters, i.e. the off-diagonal terms of Φ and Ω. In the empirical part of this paper we study the sensitivity of the estimates when $\Phi \to 0$ and $\Omega \to 0$. This is easier to do if we assume the covariance matrices to be diagonal. In any case, the results would be dominated by the diagonal elements of the matrices. We therefore assume in the remainder of the paper that $\Phi = \sigma_R^2 I$ and $\Omega = \sigma_Q^2 I$. Equations (8), (10) and (11) become, respectively,

$$b_R = (\Sigma^{-1} \otimes X'X + \frac{1}{\sigma_R^2} R'R)^{-1} (\Sigma^{-1} \otimes X')y \ , \qquad (13)$$

$$b_Q = \{ \Sigma^{-1} \otimes X'X + \frac{1}{\sigma_Q^2} (I-Q(Q'Q^{-1}Q')\}^{-1} (\Sigma^{-1} \otimes X')y \qquad (14)$$

and

$$g = \{ Q'(\Sigma \otimes (X'X)^{-1} + \sigma_Q^2 I)^{-1} Q\}^{-1} Q'\{ \Sigma \otimes (X'X)^{-1} + \sigma_Q^2 I\}^{-1} b \qquad (15)$$

In what follows we will compare the two approaches using data for U.S. manufacturing from Berndt and Wood (1975), letting σ_R and σ_Q vary and seeing how the estimated b_R, b_Q and Qg behave. Eventually, when σ_R and σ_Q approach zero both b_R and b_Q will converge to estimates with exact constraints (Brook and Wallace (1973), Ilmakunnas (1985)). In the other extreme, when σ_R and σ_Q approach infinity, b_R and b_Q converge to the unconstrained OLS estimates.

In this sense the estimators are somewhat similar to the ridge regression estimator. In ridge regression all parameters are shrunk towards zero. Here, however, not all parameters shrink since the matrices $I-Q(Q'Q)^{-1}Q'$ and $R'R$ have some zero rows. Also, the shrinking happens towards the constrained estimates, not towards zero.

Analogously to ridge regression, the use of the proposed estimators can be justified by the possibility of studying the stability of the estimates. In ridge regression this happens along the ridge trace. Here stability of b_Q, Qg and b_R when σ_Q or σ_R is changed gives an indication of how reliable any point estimates may be. However, the obtained estimates need not be superior to the unconstrained OLS estimates or the exactly constrained estimates by the mean squared error criterion, since the MSE properties of the hierarchical estimates depend on the true, but unknown parameter values (Smith (1973)); this is again similar to ridge regression (see e.g. Judge et al. (1980), ch. 12.6). Below we compare the MSE properties of the estimators in a Monte Carlo study. In any case, introducing exact or stochastic constraints leads to more efficient estimates than unconstrained OLS (see e.g. Anderson (1973)).

We can also estimate σ_Q, and Σ together with the other parameters as is done in Kiefer (1977) and Ilmakunnas (1985). In this case the estimate of σ is

$$\hat{\sigma}_Q^2 = \beta'(I-Q(Q'Q)^{-1}Q')\beta/15 \quad . \tag{16}$$

The estimate of β is

$$\hat{\Sigma} = (Y-XB)'(Y-XB)/T \tag{17}$$

where Y is a T x 3 matrix with t-th row (S_{Kt},S_{Lt},S_{Et}) and B is a 5 x 3 matrix with typical row $(\beta_{Ki},\beta_{Li},\beta_{Ei})$, i = K,L,E,M. The estimators are obtained by iterating the equation system (14), (15), (16), and (17) until convergence. The variances (16) and (17) are the modal estimates suggested by Lindley and Smith (1972), using noninformative priors for the variances.[5)] Estimating the variance of the parameters rather than fixing it a priori may make more sense when a random parameter model is used for pooling time series and cross section data, as e.g. in Swamy (1970), Smith (1973) and Trivedi (1980). However, it would be interesting to see in our simulation experiment whether we obtain an estimate of σ_Q which is close to the value which is used for generating the data.

III Results of the Monte Carlo comparison

We compared the estimators in a Monte Carlo study. The data used was that in Berndt and Wood (1975). First we obtained estimates of the parameters of the model, $\hat{\beta}$ and $\hat{\Sigma}$, using the iterative Zellner method with symmetry and homogeneity constrained. In each replication three series of normally distributed random variables were generated, which have zero mean vector and covariance matrix $\hat{\Sigma}$. A fourth series of normal random variables was used for generating random parameters, which have $\hat{\beta}$ as their mean.[6)] The error vector of the parameters is distributed as $N(0, I\sigma^2)$; σ^2 was given different values. Given the input price data, these random variables were used for generating new data of cost shares. This design corresponds to the hierarchical model. Since the parameters are stochastic, so are the constraints. As discussed above, the mixed model is more difficult to interpret as a random coefficient model.

For each value of σ^2 we made 100 replications. In each case we estimated the model with the hierarchical approach, the mixed approach, unconstrained OLS and with exact constraints, all conditionally on $\hat{\Sigma}$. We compared the estimators in terms of mean squared errors (MSE). To save space, we do not report the MSEs of all 15 parameters. In Tables 1-4 the following symbols are used: "X>>Y" implies that for all the parameters estimator Y leads to a smaller MSE than estimator X. "X>Y" implies that for all the parameters Y has smaller or equal MSE than X. "X (>) Y" denotes a case where for at most 4 parameters Y has larger MSE than X and at least for 8 parameters X has larger MSE than Y. We interpret all these three cases to give evidence that Y is in MSE sense a "better" estimator than X. Finally, in all other cases we use notation "X $\gtrless$ Y". These are interpreted as inconclusive cases.

When the true σ is zero, i.e. the parameters are nonstochastic and have exact constraints, the MSEs for the other estimators tend to be larger than for the exactly constrained estimator. The MSEs increase with the value of σ_R and σ_Q. For a given $\sigma_R = \sigma_Q$, the mixed approach tends to have smaller MSE than the first stage hierarchical estimator.

When the true σ is increased to .001 or .01, using σ_R or σ_Q equal to the true value tends to lead to a smaller MSE than other values of σ_R or σ_Q, as could be expected. However, when σ is increased to .1, using the true value does not necessarily lead to smaller MSE. Only in the case of second-stage hierarchical estimates is there a clear improvement. It can also be seen that the second-stage estimates are better than the constrained ones for a wide range of values of σ_Q when the true σ is large. This holds to some extent for the other estimators, too. For large values of σ even unconstrained OLS estimates have smaller MSE than the exactly constrained ones.

In sum, we could say that if the parameters are not stochastic, using hierarchical or mixed approach gives worse estimates, judged by MSE. If the parameters are stochastic and their variance is relatively large, there are values of σ_Q and σ_R which lead to better estimates than treating the parameters nonstochastic. Finally, for large enough true σ there is a wide range of values of σ_Q and σ_R which lead to MSE improvements over the constrained estimates.

We made some experiments with the iterative approach. σ was estimated, but Σ was fixed at $\hat{\Sigma}$, the covariance matrix of the constrained share system. Due to memory size limitations in the program used, only 25 replications were made. When $\sigma = .1$, the mean value of $\hat{\sigma}$ in the 25 replications was .06. With $\sigma = .01$ or smaller, the mean of $\hat{\sigma}$ was practically zero. Therefore it appears that there is a tendency for σ to be underestimated, which diminishes the usefulness of the iterative method.

IV An application

We used the estimators discussed above for estimating the model with the Berndt-Wood data on cost shares and input prices, conditionally on $\hat{\Sigma}$ and given values of σ_R^2 and σ_Q^2.

In tables 5 and 6 we present Allen cross elasticities of substitution. These are given by the formula $AES_{ij} = (\beta_{ij} + S_i S_j)/S_i S_j$. As cost shares we used sample mean shares. The values of σ_R^2 and σ_Q^2 used were 10^{-k} (k=0,...,8).

When k=0 and $\sigma_R^2 = \sigma_Q^2 = 1$, the elasticities were the same as those obtained from unconstrained OLS estimates. On the other hand, for the highest values of k (and smallest values of σ_R^2 and σ_Q^2) shown, the elasticities were the same as with constraints imposed exactly. For these small and large values of σ_R^2 and σ_Q^2 the mixed and hierarchical approaches lead to the same results. In the intermediate range where the most dramatic changes in the elasticities take place, the mixed estimates seem to converge more rapidly toward the constrained estimates. The main pattern of change is, however, the same. Variations in the second-stage elasticities are smoother, but some elasticities still show large changes in the same range of σ_Q^2 where the first-stage elasticities vary the most.

We can also inspect the different elasticities obtained for a given value of k. The second-stage estimates can then be used as a compromise if one needs a point estimate. Also, as shown in section III, they can lead to MSE improvement if the variance of the parameters is large. Table 6 clearly shows that in general the second-stage elasticities do not lie at the midpoint between the first-stage elasticities. Another interesting result is that in some cases the elasticities obtained cover both positive and negative values, i.e. both substitutability and complementarity of the inputs. The range of values is especially large in the case of capital and energy. This is interesting in the light of the controversy about energy-capital complementarity (e.g. Berndt and Wood (1979)). A wide variation in an elasticity when the strictness of the constraints is tightened should decrease our confidence on any particular estimate obtained since this reflects weak data information. Hence we should perhaps not reject right away the possibility of substitutability between capital and energy, since their complementarity partly results from the imposition of symmetry. For large σ_Q^2 the second-stage elasticities, too, imply energy-capital substitutability. Another interesting result is the change in the relationship of energy and materials from complementarity to substitutability when σ_Q^2 is decreased.

We also calculated own prices elasticities of demand, given by $E_i = AES_{ii} \cdot S_i$, where $AES_{ii} = (\beta_{ii} + S_i^2 - S_i)/S_i^2$. The variation in these elasticities was much less than in the Allen elasticities of substitution; therefore we report only the pair of elasticities corresponding to k=0 and k=8. In the case of E_K this was (-.17, -.39), for E_L (-.49, -.45), for E_E (-.09, -.55) and for E_M (-.13, -.22). The price elasticity of energy is hence the most sensitive to the strictness of the constraints. These ranges were the same for mixed and hierarchical models.

We also estimated the cost share system using the iterative procedure described in section II. The convergence criterion used was that changes in the elements of β were at most .0001 from previous iteration. This led to elasticities equal to those in the last columns (k=8) in Table 6, which would imply exact constraints. This result has to be taken with some reservation, since the Monte Carlo experiments in the previous section showed that the iterative method may be biased towards giving too low values of σ.

For the sake of comparison, we tested the constraints with a traditional likelihood ratio test. The test statistic, $-2\log\lambda$, where λ is the ratio of likelihood functions with and without constraints, is 11.68. This is above the critical value, 10.6, of the χ^2 distribution with 6 degrees of freedom at the 10 percent significance level but below the critical value at the 5 percent level, 12.6. Hence the constraints are only marginally accepted.[7)] Studying the sensitivity of the estimates to the imposition of the constraints seems therefore worthwhile.

IV Conclusions

The purpose of this paper has been to compare different approaches to imposing parameter constraints stochastically in a cost function model. Monte Carlo evidence suggests that if the parameter constraints are indeed stochastic with a large enough variance, mean squared error can be decreased by adopting a mixed or a hierarchical approach to estimating the model. In an application with the Berndt-Wood (1975) data it appears that capital-energy substitutability is clearly a possibility if the parameters and hence the constraints are treated as stochastic. In contrast, when the constraints are fixed, energy-capital complementarity is found.

Given different prior strictness of the parameter constraints we obtain a range of elasticities. It seems useful to report the whole range so that the reader can judge whether the point estimates are too sensitive to the imposition of the constraints to be useful e.g. in making forecasts of industrial demand for energy or of substitution between inputs following changes in relative input prices.

Acknowledgement

The first version of this paper was written when I was a research fellow at the International Institute of Management, Wissenschaftszentrum Berlin. I am thankful to a referee for useful comments.

Table 1: MSE comparison, $\sigma = 0$

	$\gtrless$	Mixed, $\sigma_R = .001$	<<	Mixed, $\sigma_R = .01$	<<	Mixed, $\sigma_R = .1$	<<	
		(<)		<		<		
CJGLS	(<)	1st stage, $\sigma_Q = .001$	<<	1st stage, $\sigma_Q = .01$	<<	1st stage, $\sigma_Q = .1$	<<	OLS >> CJGLS
		$\gtrless$		$\gtrless$		$\gtrless$		
	(<)	2nd stage, $\sigma_Q = .001$	<<	2nd stage, $\sigma_Q = .01$	<<	2nd stage, $\sigma_Q = .1$	$\gtrless$	

Note:

Mixed	=	mixed regression estimates
1st stage	=	hierarchical model, 1st stage estimates
2nd stage	=	hierarchical model, 2nd stage estimates
OLS	=	ordinary least squares
CJGLS	=	constrained joint generalized least squares
>>	=	all parameter estimates have larger MSE
>	=	all parameter estimates have larger or equal MSE
(>)	=	for at most 4 parameter estimates MSE is smaller and for at least 8 parameter estimates MSE is larger
$\gtrless$	=	inconclusive

Table 2: MSE comparison, $\sigma = .001$

	$\gtrless$	Mixed, $\sigma_R = .001$	$(<)$	Mixed, $\sigma_R = .01$	$<<$	Mixed, $\sigma_R = .1$	$<<$	
		$\gtrless$		$<<$		$(<)$		
CJGLS	$\gtrless$	1st stage, $\sigma_Q = .001$	$<<$	1st stage, $\sigma_Q = .01$	$<<$	1st stage, $\sigma_Q = .1$	$<<$	OLS >> CJGLS
		$\gtrless$		$\gtrless$		$\gtrless$		
	$\gtrless$	2nd stage, $\sigma_Q = .001$	$<<$	2nd stage, $\sigma_Q = .01$	$<<$	2nd stage, $\sigma_Q = .1$	$\gtrless$	

Note: see Table 1

Table 3: MSE comparison, $\sigma = .01$

	>	Mixed, $\sigma_R = .001$	(>)	Mixed, $\sigma_R = .01$	(<)	Mixed, $\sigma_R = .1$	<	
		>>		$\gtrless$		<		
CJGLS	(>)	1st stage, $\sigma_Q = .001$	(>)	1st stage, $\sigma_Q = .01$	(<)	1st stage, $\sigma_Q = .1$	<	OLS $\gtrless$ CJGLS
		$\gtrless$		(>)		(>)		
	>>	2nd stage, $\sigma_Q = .001$	>>	2nd stage, $\sigma_Q = .01$	(<)	2nd stage, $\sigma_Q = .1$	(<)	

Note: see Table 1

Table 4: MSE comparison, $\sigma = .1$

	>>	Mixed, $\sigma_R = .001$	(>)	Mixed, $\sigma_R = .01$	$\gtrless$	Mixed, $\sigma_R = .1$	$\gtrless$	
		(>)		(>)		$\gtrless$		
CJGLS	>>	1st stage, $\sigma_Q = .001$	(>)	1st stage, $\sigma_Q = .01$	$\gtrless$	1st stage, $\sigma_Q = .1$	(<)	OLS (<) CJGLS
		$\gtrless$		(>)		>		
	>>	2nd stage, $\sigma_Q = .001$	>>	2nd stage, $\sigma_Q = .01$	>>	2nd stage, $\sigma_Q = .1$	(<)	

Note: see Table 1

Table 5: Allen Elasticities of Substitution: The Mixed Approach

	k=0	1	2	3	4	5	6	7	8
AES_{KL}	-.69	-.67	-.47	.25	.80	.93	.97	.97	.97
AES_{LK}	2.34	2.32	2.17	1.63	1.17	.98	.97	.97	.97
AES_{KE}	17.40	17.23	15.77	9.14	.35	-2.77	-3.11	-3.14	-3.14
AES_{EK}	-1.44	-1.47	-1.68	-2.52	-3.23	-3.23	-3.15	-3.14	-3.14
AES_{KM}	.57	.56	.48	.25	.34	.44	.43	.43	.43
AES_{MK}	-.62	-.60	-.49	-.08	.30	.43	.43	.43	.43
AES_{LE}	6.46	6.39	5.76	3.31	1.34	.79	.66	.65	.64
AES_{EL}	.22	.23	.35	.77	.95	.75	.66	.65	.64
AES_{LM}	.28	.27	.26	.29	.48	.57	.58	.58	.58
AES_{ML}	.81	.81	.76	.60	.56	.57	.58	.58	.58
AES_{EM}	.57	.56	.51	.38	.52	.77	.84	.85	.85
AES_{ME}	-4.03	-3.98	-3.53	-1.66	.23	.78	.85	.85	.85

Note: $k = \log_{10}(1/\sigma_R^2)$

Table 6: Allen Elasticities of Substitution: The Hierarchical Approach

	k=0	1	2	3	4	5	6	7	8
First stage									
AES_{KL}	-.69	-.68	-.58	.01	.72	.90	.96	.97	.97
AES_{LK}	2.34	2.33	2.24	1.74	1.17	.95	.96	.97	.97
AES_{KE}	17.40	17.30	16.36	10.94	2.39	-1.99	-2.99	-3.12	-3.14
AES_{EK}	-1.44	-1.46	-1.60	-2.38	-3.33	-3.45	-3.21	-3.15	-3.14
AES_{KM}	.57	.57	.53	.37	.31	.44	.44	.43	.43
AES_{MK}	-.62	-.61	-.54	-.17	.29	.45	.44	.43	.43
AES_{LE}	6.47	6.42	6.01	3.81	1.52	1.04	.73	.65	.65
AES_{EL}	.22	.22	.29	.65	1.03	.92	.70	.65	.65
AES_{LM}	.28	.28	.29	.36	.49	.55	.58	.58	.58
AES_{ML}	.82	.81	.79	.67	.56	.56	.58	.58	.58
AES_{EM}	.57	.56	.54	.44	.41	.63	.81	.85	.85
AES_{ME}	-4.04	-4.00	-3.71	-2.08	-.06	.62	.82	.85	.85
Second stage									
AES_{KL}	.97	.97	.98	1.00	.98	.92	.95	.97	.97
AES_{KE}	.35	.35	.33	.16	-1.10	-2.66	-3.08	-3.13	-3.14
AES_{KM}	.20	.20	.19	.17	.26	.43	.44	.43	.43
AES_{LE}	2.99	2.98	2.84	2.11	1.33	1.02	.72	.65	.65
AES_{LM}	.37	.37	.38	.42	.50	.55	.58	.58	.58
AES_{MK}	-.29	-.29	-.25	-.01	.36	.66	.82	.85	.85

Note: $k = \log_{10}(1/\sigma_Q^2)$

Footnotes

1) Pekka Ilmakunnas, Research Institute of the Finnish Economy, Lönnrotinkatu 4 B, SF-00120 Helsinki, Finland.

2) See discussion in Kiefer (1977) and the references cited there. One should also note that testing the constraints implied by optimizing behavior means testing the behavioral assumption. However, an alternative theory that includes absence of the constraints may not exist; see Phlips (1976).

3) Leamer and Leonard (1983) recommend analyzing the fragility of estimates, but use a different framework.

4) See the series of papers by Laitinen (1978), Meisner (1979), Bera, Byron and Jarque (1981), Fiebig and Theil (1983) and Theil and Rosalsky (1984).

5) The divisor in (17) depends on whether one considers the joint or marginal modes of the parameters (see O'Hagan (1976)). We have used the marginal modes and hence the divisor T, since this corresponds more closely to maximum likelihood estimation.

6) The errors were generated using the random number generator in TSP, version 4.0. The error terms for the share equations were obtained as follows. Let ε_i, i=1,2,3, be the generated N(0,1) error terms and $\varepsilon = [\varepsilon_1, \varepsilon_2, \varepsilon_3]$ so that $E(\varepsilon'\varepsilon) = I$. We decomposed the covariance matrix to $\hat{\Sigma} = P'P$ where P is an upper triangular matrix. Postmultiply ε by P to obtain new errors $\eta = \varepsilon P$. It is easy to see that $E(\eta'\eta) = \hat{\Sigma}$ and $vec(\eta) \sim N(0, \hat{\Sigma} \otimes I)$.

7) The references cited in footnote 4 may imply that since there is a tendency for the constraints to be rejected, the marginal acceptance of the constraints in our case might actually give fairly strong evidence for the constraints. However, in our case the equation system is small and hence the bias towards rejection should not bee too serious.

References

Anderson, R.J., "A Nesting Theorem on Efficiency of Bayesian or Mixed Regression Model Estimators", International Economic Review 14, 1973, 789-790.

Bera, A.K., R.P. Byron and C.M. Jarque, "Further Evidence on Asymptotic Tests for Homogeneity and Symmetry in Large Demand Systems", Economics Letters 8, 1981, 101-105.

Berndt, E.R. and D.O. Wood, "Technology, Prices and the Derived Demand for Energy", Review of Economics and Statistics 57, 1975, 259-268.

Berndt, E.R. and D.O. Wood, "Engineering and Econometric Interpretations of Energy-Capital Complementarity", American Economic Review 69, 1979, 342-354.

Brook, R. and T.D. Wallace, "A Note on extraneous Information in Regression", Journal of Econometrics 1, 1973, 315-316.

Evans, G. and K.D. Patterson, "The Lagrange Multiplier Test for Autocorrelation in the Presence of Linear Restrictions", Economics Letters 17, 1985, 237-241.

Fiebig, D.G. and H. Theil, "The Two Perils of Symmetry-Constrained Estimation of Demand Systems", Economics Letters 13, 1983, 105-111.

Haitovsky, Y., "The linear Hierarchical Model: I. Estimation", MRG Working Paper 7902, University of Southern California, 1979.

Ilmakunnas, P., "Bayesian Estimation of Cost Functions with Stochastic or Exact Constraints on Parameters", International Economic Review 26, 1985, 111-134.

Judge, G.G., W.E. Griffiths, R.C. Hill and T.-C. Lee, The Theory and Practice of Econometrics, New York: Wiley, 1980.

Kiefer, N.M., "A Bayesian Analysis of Commodity Demand and Labor Supply", International Economic Review 18, 1977, 209-218.

Laitinen, K., "Why Is Demand Homogeneity So Often Rejected?", Economics Letters 1, 187-191.

Leamer, E. and H. Leonard, "Reporting the Fragility of Regression Estimates", Review of Economics and Statistics 65, 1983, 306-317.

Lindley, D.V. and A.F.M. Smith, "Bayes Estimates for the Linear Model", Journal of Royal Statistical Society B 34, 1972, 1-18.

Meisner, J.F., "The Sad Fate of the Asymptotic Slutsky Symmetry Test for Large Systems", Economic Letters 2, 1979, 231-233.

O'Hagan, A., "On Posterior Joint and Marginal Modes", Biometrika 63, 1976, 329-333.

Paulus, J.D., "Mixed Estimation of a Complete System of Consumer Demand Equations", Annals of Economic and Social Measurement 4, 1975, 117-131.

Phlips, L., "Comment", Annals of Economic and Social Measurement 4, 1975, 45-48.

Smith, A.F.M., "A General Bayesian Linear Model", Journal of Royal Statistical Society B 35, 1973, 67-75.

Swamy, P.A.V.B., "Efficient Inference in a Random Coefficient Regression Model", Econometrica 38, 1970, 311-323.

Swamy, P.A.V.B. and J.S. Mehta, "Ridge Regression Estimation of the Rotterdam Model", Journal of Econometrics 22, 1983, 365-390.

Theil, H., Principles of Econometrics, New York: Wiley, 1971.

Theil, H. and A.S. Goldberger, "On Pure and Mixed Statistical Estimation in Economics", International Economic Review 2,1961, 65-78.

Trivedi, P.K., "Small Samples and Collateral Information: An Application of the Hyperparameter Model", Journal of Econometrics 12, 1980, 301-318.

Tsurumi, H., H. Wago and P. Ilmakunnas, "Gradual Switching Multivariate Regression Models with Stochastic Cross-Equational Constraints and an Application to the KLEM Translog Production Model", 1985, forthcoming in Journal of Econometrics.

Zellner, A., "The Bayesian Approach and Alternatives in Econometrics", in S.E. Fienberg and A. Zellner, eds., Studies in Bayesian Econometrics and Statistics, Amsterdam: North Holland, 1975, 39-54.

ELINKEINOELÄMÄN TUTKIMUSLAITOS (ETLA)
The Research Institute of the Finnish Economy
Lönnrotinkatu 4 B, SF-00120 HELSINKI 12 Puh./Tel. (90) 601 322

KESKUSTELUAIHEITA - DISCUSSION PAPERS ISSN 0781-6847

No 170 ROBERT A. HART, Wage Supplements through Collective Agreement or Statutory Requirement? 04.02.1985. 26 p.

No 171 AKINARI HORII, Financial Innovation and its Implications for the Japanese Financial System. 18.03.1985. 8 p.

No 172 PAUL D. McNELIS, Indexing and Macroeconomics: A Survey of Theoretical Developments and International Experience. 02.04.1985. 44 p.

No 173 EERO PYLKKÄNEN - PENTTI VARTIA, Some Comments on Fine-Tuning Macro-Model Forecasts. 29.05.1985. 22 p.

No 174 CHRISTIAN EDGREN, Automatic versus Discretionary Tax-Policy Effects Revisited. 05.06.1985. 44 p.

No 175 TAUNO KALLINEN, Työn tuottavuuden taso ja ansiotaso Suomen, Ruotsin, Norjan ja Tanskan tehdasteollisuudessa vuosina 1973 ja 1982. 11.06.1985. 44 s.

No 176 J. LESKELÄ - J. MOILANEN, A Comparison of Recent Medium-Term Forecasts. 17.06.1985. 34 p.

No 177 PAAVO SUNI, Efektiivisen nimellisen valuuttakurssin mittaaminen ja mittarien tulkinta. 20.06.1985. 42 s.

No 178 MARKKU LAMMI, Alkoholi- ja polttoaineverojen kertymästä. 27.09.1985. 49 s.

No 179 YRJÖ O. VARTIA, Tunnetuissa pisteissä epäsäännöllisen funktion numeerinen integrointi. 21.10.1985. 8 s.

No 180 ERKKI KOSKELA, On Monopolistic Tax Evasion under Direct and Indirect Taxation. 23.10.1985. 17 p.

No 181 PEKKA ILMAKUNNAS, Stochastic Constraints on Cost Function Parameters: Mixed and Hierarchical Approaches. 31.10.1985. 26 p.

No 182 HIROKI TSURUMI - HAJIME WAGO - PEKKA ILMAKUNNAS, Gradual Switching Multivariate Regression Models with Stochastic Cross-Equational Constraints and an Application to the Klem Translog Production Model. 01.11.1985. 28 p.

Elinkeinoelämän Tutkimuslaitoksen julkaisemat "Keskusteluaiheet" ovat raportteja alustavista tutkimustuloksista ja väliraportteja tekeillä olevista tutkimuksista. Tässä sarjassa julkaistuja monisteita on rajoitetusti saatavissa ETLAn kirjastosta tai ao. tutkijalta.

Papers in this series are reports on preliminary research results and on studies in progress; they can be obtained, on request, by the author's permission.

0033A/01.11.1985