Volume Title: Annals of Economic and Social Measurement, Volume 4, number 1

Volume Author/Editor: Sanford V. Berg, editor

Volume Publisher: NBER

Volume URL: http://www.nber.org/books/aesm75-1

Publication Date: 1975

Chapter Title: Mixed Estimation of a Complete System of Consumer Demand Equations

Chapter Author: John D. Paulus

Chapter URL: http://www.nber.org/chapters/c10220

Chapter pages in book: (p. 117 - 131)

Annals of Economic and Social Measurement, 4/1, 1975.

MIXED ESTIMATION OF A COMPLETE SYSTEM OF CONSUMER DEMAND EQUATIONS

BY JOHN D. PAULUS*

The Theil-Goldberger "mixed estimator" (I.E.R. 1961) is used to estimate the parameters of a complete system of consumer demand equations. The mixed estimator combines the sample evidence, which consists of observations on the consumption of each of 14 goods in the Netherlands from 1922 to 1963, with stochastic prior estimates of the income elasticities of all goods (complete ignorance is assumed regarding price elasticities). This provides a simple yet flexible method of reducing multicollinearity among the predetermined price and income variables. A statistical test of the compatibility of the prior and sample information is carried out and measures of the posterior precision of the point estimates attributable to the two information sources are computed. It is shown that the use of stochastic restrictions against the income coefficients reduces the "effective number of unconstrained parameters" in the demand model by about one-third.

1. INTRODUCTION

In this study the parameters of a complete system of consumer demand equations will be estimated using Barten's [2] Dutch data on 14 commodity groups. Demand systems involving large numbers of commodities, say 10 or more, are usually characterized by multicollinearity among the independent variables. See, for example, the recent empirical studies of complete demand systems by Barten [4] and Byron [5].

In order to avoid the multicollinearity problem two types of constraints will be imposed on the parameters of the demand model. First, it is assumed that the utility function is strongly separable. This enables us to eliminate many price coefficients from the model. Secondly, the model is further restricted by imposing "probabilistic" constraints on a subset of the parameter vector. Specifically, we utilize subjective prior estimates of the income elasticities of the 14 commodities together with a covariance matrix of these estimates to constrain the income coefficients of the model. The statistical device to be used in combining the prior and sample information is the Theil-Goldberger [27] "mixed estimator." It will be shown that the imposition of probabilistic constraints reduces the "effective number of unconstrained parameters" in the model by about 35 percent.

2. MIXED AND BAYESIAN METHODS

If certain axioms on rational behavior under uncertainty are accepted, it can be shown that a decision maker will act as though he has *assigned prior probabilities* to the various outcomes of an uncertain event, and was trying to maximize expected utility.[1] The existence of these prior probabilities in a decision theoretic model

* This work was undertaken as part of a doctoral dissertation at the University of Chicago. I am indebted to H. Theil for his careful supervision of that work and for making many helpful comments on an earlier draft of this paper. I also wish to thank P. A. V. B. Swamy for his assistance in pointing out and explaining the fiducial probability interpretation of the mixed estimator.

[1] See Ferguson [8], pp. 11–21.

supports the Bayesian position in econometrics that a joint prior distribution on the relevant parameter vector always exists. Consider the linear stochastic model

$$y = X\beta + \varepsilon \qquad E\varepsilon = 0 \qquad V(\varepsilon) = \Omega \tag{1}$$

where y is a T-vector of observations on the dependent variable, X is a T by k matrix of fixed independent variables, ε is a T-vector of random errors, β is a random k-vector of parameters, and Ω is a known T by T covariance matrix of ε. The Bayesian analyst combines his prior *pdf* on β, $p(\beta)$, with the likelihood function to produce the posterior distribution of β,

$$p(\beta|y, X, \Omega) \propto p(\beta) l(y|X, \beta, \Omega).$$

This distribution contains all the information available on β. If a point estimate of β is desired, the Bayesian selects that estimate that minimizes the expected value of a loss function defined over β.

The major objection to the Bayesian model concerns the difficulty of defining an operational prior *pdf* over β, especially when the number of parameters is large. This problem is widely recognized by Bayesian and classical econometricians alike.[2] A simple alternative to working with a complete prior *pdf* is to utilize only the first two moments of the prior. Chipman [6], pp. 1104–1106 considers the problem of finding a linear minimum mean squared error (MMSE) estimator of the random parameter vector β in (1) when prior information of the form

$$E\beta = \bar{\beta} \qquad V(\beta) = E(\beta - \bar{\beta})(\beta - \bar{\beta})' = V$$

is available. Formally the problem is to minimize the "risk matrix," $E(\hat{\beta} - \beta) \cdot (\hat{\beta} - \beta)'$, subject to $\hat{\beta} = a + By$, where a and B are to be determined. The problem is analogous to the Bayesian point estimation problem of finding a measure of central tendency that minimizes expected loss. Chipman's MMSE estimator is

$$\begin{aligned}\hat{\beta} &= \bar{\beta} + (X'\Omega^{-1}X + V^{-1})^{-1}X'\Omega^{-1}(y - X\bar{\beta}) \\ &= (X'\Omega^{-1}X + V^{-1})^{-1}X'\Omega^{-1}y + (I - (X'\Omega^{-1}X + V^{-1})^{-1}X'\Omega^{-1}X)\bar{\beta} \\ &= (X'\Omega^{-1}X + V^{-1})^{-1}(X'\Omega^{-1}y + V^{-1}\bar{\beta})\end{aligned} \tag{2}$$

where, as in the full Bayesian model, y and $\bar{\beta}$ are interpreted as fixed observations.[3]

The Theil–Goldberger [27] mixed estimator is the classical analogue of Chipman's MMSE estimator. See Swamy and Mehta [19] and Mehta and Swamy [12] for an analysis of the finite sample distribution of the mixed estimator and other properties. The mixed estimator is developed from the classical linear model

$$y = X\beta + \varepsilon \qquad E\varepsilon = 0 \qquad V(\varepsilon) = \Omega \tag{3.1}$$

where β is taken as a *fixed* k-vector of parameters. Now, suppose stochastic prior information is available on a subset of the parameter vector

$$r = R\beta + v \qquad Ev = 0 \qquad V(v) = V \tag{3.2}$$

[2] See e.g., G. Kaufman [11], p. 206. Also see T. J. Rothenberg [18] and A. Zellner [30], p. 239 for remarks on problems relating to the use of natural conjugate priors in the multivariate regression model.

[3] The estimator (2) was developed independently by Rao [17], p. 192. See also Ericson [7].

where r is a q-vector of random prior estimates of $R\beta$, R is a $q \times k$ $(q \leq k)$ matrix[4] of full row rank and v is a q-vector of random errors. Combining (3.1) ad (3.2) and applying GLS yields the mixed estimator

$$\hat{\beta}_M = (X'\Omega^{-1}X + R'V^{-1}R)^{-1}(X'\Omega^{-1}y + R'V^{-1}r) \tag{4}$$

with covariance matrix

$$V(\hat{\beta}_M) = (X'\Omega^{-1}X + R'V^{-1}R)^{-1}. \tag{5}$$

The estimator (4) is the MVLU of β and is based on the assumption that the prior information is uncorrelated with the sample evidence.[5] When $\hat{\beta}_M$ is computed, both r and y are regarded as realizations of random variables.

When $R = I$ (prior assessments are available on all parameters), the mixed estimator is algebraically (and hence computationally) equivalent to Chipman's MMSE estimator. Conceptually the two estimators differ because β and y are considered to be fixed in (2), while r and y are random in (4).

Frazer's [9, 10] frequency interpretation of fiducial probability may be used to justify the interpretation of r as a random variable in (4). (See also Mehta and Swamy [14] and Swamy and Mehta [20]). For example, suppose we have a random variable $\tilde{r} = \beta - \tilde{v}$, $E\tilde{v} = 0$, $V(\tilde{v}) < \infty$, which represents data based prior information on the parameter β. Now consider a drawing from this distribution yielding the fixed value r. Using Frazer's linear translation rules, we can define the fiducial probability distribution of β in terms of the random variable $\tilde{\beta} = r + \tilde{v} = \beta + (r - \beta) + \tilde{v} = \beta + \tilde{u}$, $E\tilde{u} = r - \beta \neq 0$. This equation corresponds to equation (3.2) when $R = I$. Note that the assumption of unbiased prior information is dropped. We therefore consider the implications of relaxing the requirement, $Ev = 0$ in (3.2).

Consider the system (3.1) and (3.2) with $Ev = d \neq 0$ replacing $Ev = 0$ in (3.2). The covariance matrix of v will be $E(v - d)(v - d)' = Evv' - (Ed)(Ed)' = V_0 - dd' = V$, so that $Evv' = V_0 = V + dd'$. The mixed estimator (4) is biased when $Ev = d$ and the second moment matrix of $\hat{\beta}_M$ around β is

$$\begin{aligned} E(\hat{\beta}_M - \beta)(\hat{\beta}_M - \beta)' &= (X'\Omega^{-1}X + R'V^{-1}R)^{-1} \\ &+ (X'\Omega^{-1}X + R'V^{-1}R)^{-1}(R'V^{-1}dd'V^{-1}R)(X'\Omega^{-1}X + R'V^{-1}R)^{-1}. \end{aligned} \tag{6}$$

From (6) it is seen that the true second moment matrix around β is equal to $V(\hat{\beta}_M|d = 0)$ plus the positive semidefinite matrix on the second line of (6). Computed standard errors based on $(X'\Omega^{-1}X + R'V^{-1}R)^{-1}$ will therefore understate the true variability of the mixed point estimates when $d \neq 0$.

Alternatively, suppose we compute $\hat{\beta}_M = (X'\Omega^{-1}X + R'V_0^{-1}R)^{-1}(X'\Omega^{-1}y + R'V_0^{-1}r)$ where $V_0 = Evv'$. Theil [24], p. 352, (prob 8.3) points out that the matrix $(X'\Omega^{-1}X + R'V_0^{-1}R)^{-1}$ is then the matrix of second moments of $\hat{\beta}_M$ around β. The mixed estimator, though still biased, is then a better estimator of β in a mean squared error sense than the sample GLS estimator, $\hat{\beta} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y$,

[4] The usual form of R is $[I_q : 0]$ which indicates that prior assessments are entered against the first q parameters.

[5] H. Theil [26] has recently developed an extended mixed estimator that allows the prior information to be correlated with the sample.

since $V(\hat{\beta}) = (X'\Omega^{-1}X)^{-1}$ exceeds $(X'\Omega^{-1}X + R'V_0^{-1}R)^{-1}$ by a positive semi-definite matrix. The requirement $Ev = 0$ in (3.2) is therefore unnecessary provided the second moment matrix around zero (rather than the covariance matrix) of the prior error terms is used in the computation of the mixed estimator.

3. Choosing an Estimator

The mixed estimator provides a flexible method for introducing stochastic prior information against a subset of the parameter vector. In contrast, Chipman's MMSE estimator requires that prior information be available on every element of the parameter vector,[6] while the Bayesian has to resort to "mixed" prior *pdfs*—partly proper and partly diffuse—when informative prior information is available on only a proper subset of the parameter vector. The number of parameters to be directly estimated in this study is 17. It would be a most difficult problem to define a joint prior *pdf*, even over a subset of these parameters, which satisfies the various constraints imposed on the parameters of the model by economic theory (see Section 4). It does, however, seem reasonable to work with the first and second moments of a subset of the parameter vector. The mixed estimator will therefore be used in Section 6 to obtain estimates of the demand parameters. Other important, though less compelling reasons for preferring the mixed procedure to the Bayesian model in this study are given below.

There is some evidence that the prior distributions or odds developed by empirical analysts are often not, in Raiffa's terminology, empirically validated. That is, they conflict with accurate post-sample measurements of the unknown parameters. In cases where the true values of the parameters are not revealed by the sample, as in econometric model building and estimation, it would be desirable to have a statistical test of the compatibility of the prior and sample information.

Theil [21] develops such a test within the framework of the mixed estimation problem. The hypothesis to be considered is

$$H_0 : E(r - R\hat{\beta}) = 0$$

where $\hat{\beta} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y$ is the GLS estimator of the model (3.1) based on the sample information. In words, the test considers the equality of the expectations of the prior and sample estimates. The test statistic is

$$(r - R\hat{\beta})'(R(X'\Omega^{-1}X)^{-1}R' + V_0)^{-1}(r - R\hat{\beta}) \tag{7.1}$$

which is distributed $\chi^2(q)$ under H_0 when the sample disturbances and the prior errors are normally distributed. The number of degrees of freedom, q, is equal to the number of elements in r (the number of linearly independent prior assessments of β).[7]

If H_0 is accepted, it would then be interesting to determine what share of the precision of the mixed estimates is contributed by the prior estimates. Theil [21]

[6] This objection can be weakened by letting the variance of a parameter about which we are uninformed grow arbitrarily large. This diminishes the weight given to the unknown prior mean in (2).

[7] See Mehta and Swamy [13] for the finite sample distribution of the compatibility statistic.

developed such a measure based on four rather simple axioms. His prior and sample precision shares are

$$\alpha_p = \frac{1}{k}\text{tr } R'V_0^{-1}R(X'\Omega^{-1}X + R'V_0^{-1}R)^{-1} \tag{7.2}$$

$$\alpha_s = \frac{1}{k}\text{tr } X'\Omega^{-1}X(X'\Omega^{-1}X + R'V_0^{-1}R)^{-1}$$

where k is the number of parameters being estimated and tr is the "trace" operator. It is easily seen that if the elements of V grow without bound, α_p converges to zero and α_s to 1. Conversely, if Ω increases indefinitely, the sample share will converge to zero and the prior share to one. Notice also that $\alpha_s + \alpha_p = 1$.

In Section 7 the sample share of the posterior precision will be used to derive a measure of the "effective number of unconstrained parameters" in the model after probabilistic constraints are imposed on the income coefficients. This will tell us how effective the use of stochastic prior information has been in reducing the multicollinearity problem.

4. The Demand Model

The Rotterdam model in finite changes is[8]

$$w_{it}^*Dq_{it} = \mu_i Dq_t + \sum_j v_{ij}(Dp_{jt} - \sum_k \mu_k Dp_{kt}) + \varepsilon_{it} \qquad i = 1, \ldots n \tag{8}$$

where

$$Dq_{it} = \Delta(\log q_{it}) = \log q_{it} - \log q_{i,t-1}$$

$$Dp_{it} = \Delta(\log p_{it}) = \log p_{it} - \log p_{i,t-1}$$

$$w_{it}^* = \tfrac{1}{2}(w_{i,t-1} + w_{it}) \quad \text{with } w_{it} = p_{it}q_{it}/m_t$$

$$Dq_t = \sum_k w_{kt}^* Dq_{kt}$$

$$\varepsilon_{it} = \text{random disturbance term}$$

with p_{it} and q_{it} being, respectively, the i-th price and quantity at t, and m_t is income at t. Dq_t can be shown to be a very close approximation to the log change in the true index of real income evaluated at prices equal to the geometric mean of those prevailing in $t-1$ and t. The parameters of the model are μ_i, the i-th marginal budget share (i.e., the proportion of an increment to the consumer's budget which is allocated to good i), and v_{ij}, the coefficient of the j-th deflated price in the i-th demand equation.[9] It can be shown that it is natural to define specific substitution and complementarity in terms of the sign of v_{ij} as follows:

$v_{ij} > 0$ goods i and j are specific substitutes

$v_{ij} < 0$ goods i and j are specific complements

[8] See Barten [1, 3, 4] and Theil [22, 23, 24, 25] for a discussion of the Rotterdam model and its statistical implementation.

[9] The parameters of (8) are, in principle, not constant, but functions of income and prices. When (8) is implemented, μ_i and v_{ij} are treated as constants to economize on the number of parameters and to make the estimation procedure more tractable. This amounts to a linearization of the demand equation (linear in $Dq_t, Dp_{1t}, \cdots, Dp_{nt}$).

The parameters are subject to the following constraints:

$$\sum_{i}^{n} \mu_i = 1 \tag{9.1}$$

$$\sum_{j}^{n} \nu_{ij} = \varphi \mu_i \tag{9.2}$$

$$\nu_{ij} = \nu_{ji} \tag{9.3}$$

where φ is the so-called "income flexibility."[10] These constraints will be substituted into (8) in Section 6 to eliminate many parameters from the model and to simplify the estimation procedure.

5. Data and Prior Judgements on the Income Coefficients

The data in this study consist of annual observations on the consumption of goods and services in the Netherlands from 1922 to 1963.[11] The 14 commodities are listed in column 1 of Table 1, while their average budget shares for the 29 year sample period are given in column 2. Notice that the goods are partitioned into three basic groups.

Prior estimates of the income elasticities of the 14 goods are given in column 3 with "standard deviations" in parentheses.[12] We confine ourselves to prior assessments of the income coefficients of the model because there have been numerous studies on the income sensitivity of the demand for goods and services that form the basis for at least a rough approximation to the true values of the unknown elasticities.[13] In contrast, much uncertainty remains regarding price sensitivity, especially when interpreted as the effect of a price change compensated by an income change that restores the marginal utility of income (which is how the ν_{ij} coefficients are to be interpreted).

The prior estimates of the first 13 elasticities are obtained by making pairwise comparisons of the luxury-necessity characteristics of goods in the same group (see the grouping in Table 1). For example, the prior estimate of the income elasticity of Fish is 0.8 which is halfway between the prior estimates for Dairy

[10] Formally, $\varphi = [\partial \log \lambda / \partial \log m]^{-1}$ where λ is the marginal utility of income. The income flexibility is therefore equal to the reciprocal of the income elasticity of the marginal utility of income.

[11] The data were collected by Barten [2]. The World War II transition period is omitted, leaving 29 observations on each of the 14 commodities. See H. Theil [23], Ch. 5 for a reproduction of the relevant data. The data for the 14th commodity are derived from House rent (subscript H) and Services and other commodities (subscript S) as follows

$$w^*_{14,t} Dq_{14,t} = w^*_{Ht} Dq_{Ht} + w^*_{St} Dq_{St}$$

$$Dp_{14,t} = (w^*_{Ht} Dp_{Ht} + w^*_{St} Dp_{St})/(w^*_{Ht} + w^*_{St})$$

where $w^*_{14,t} = (w_{14,t-1} + w_{14,t})/2$ with $w_{14,t} = w_{Ht} + w_{St}$. The data for $i = 10$ are similarly derived from the original data on Household durables and Other durables.

[12] These "standard deviations" are chosen conservatively to reflect a sensible second moment of each estimate around the true parameter. We shall continue to use "standard deviation" in the text to refer to these second moments.

[13] See, for example, Prais and Houthakker [16] and Wold [28]. H. Theil provided considerable insight into Dutch consumption habits, so that the estimates are partly subjective.

TABLE 1

AVERAGE VALUE SHARES AND PRIOR ESTIMATES OF MARGINAL SHARES FOR 14 COMMODITY GROUPS IN THE NETHERLANDS, 1922-63

Commodity Group (1)	Average Value Share* (2)	Prior Estimate and Standard Deviation of Income Electricity (3)	Marginal Share (4)
Food			
1. Bread	0.0352	0.2 (0.10)	0.0070 (0.0035)
2. Groceries	0.0565	0.4 (0.15)	0.0227 (0.0085)
3. Dairy Products	0.0723	0.6 (0.15)	0.0434 (0.0108)
4. Vegetables and fruit	0.0443	0.6 (0.20)	0.0265 (0.0088)
5. Meat	0.0760	1.0 (0.20)	0.0760 (0.0152)
6. Fish	0.0071	0.8 (0.25)	0.0057 (0.0018)
7. Pastry, chocolate, ice cream	0.0312	0.8 (0.25)	0.0249 (0.0078)
Beverages and Tobacco			
8. Beverages	0.0271	1.5 (0.30)	0.0406 (0.0081)
9. Tobacco	0.0375	0.6 (0.25)	0.0225 (0.0094)
Durables/Remainder			
10. Household and other durables	0.0941	2.0 (0.50)	0.1882 (0.0471)
11. Water, light, and heat	0.0549	0.8 (0.30)	0.0439 (0.0165)
12. Clothing and other textiles	0.1328	2.0 (0.40)	0.2656 (0.0531)
13. Footwear	0.0143	1.5 (0.30)	0.0215 (0.0043)
14. Other goods and services	0.3167	0.7 (0.35)	0.2115 (0.1100)

* Share of total *per capita* expenditure in each year, averaged over all years of the sample.

products and Meat. It was felt that, though Fish has much in common with Meat, it probably has a lower income elasticity because the consumption of some Fish (such as Herring) is similar to certain Dairy products, such as cheese and eggs, which are often eaten with bread and crackers.

The income elasticity of the 14th commodity is difficult to appraise because of its heterogeneous content. However, it appears in Section 6 that prior estimates of only the first 13 marginal shares are needed to compute mixed estimates, so that it is unnecessary to introduce a prior estimate of the 14th income elasticity. An estimate is given in Table 1 for the sake of completeness. It is based on the identity $\sum_i w_i\sigma_i = 1$ where σ_i is the income elasticity of the i-th good.

The standard deviation of a point estimate reflects the confidence we have in the quality of the estimate. Those commodities with a more heterogeneous composition are assigned larger standard deviations because it is felt that it is more difficult to assess the income sensitivity of such goods. Also, those goods with larger income elasticities are assigned larger standard deviations, reflecting a greater uncertainty about the income elasticity in absolute terms, though not necessarily in relative terms.

If we multiply the prior estimates of the income elasticities by the budget shares, we obtain prior estimates of the marginal budget shares. Similarly, by multiplying the standard deviations of the prior estimates of the income elasticities by the budget shares, we obtain the standard deviations of the prior estimates of the marginal budget shares.[14] These are shown in column 4 of

[14] This follows from treating the budget shares as constants.

Table 1 where the average budget shares in column 2 are used to approximate the budget shares.

The implementation of the mixed estimation method requires a specification of the covariances of the prior estimates. The covariances should reflect the fact that the prior estimates of the first 13 income elasticities were based on pairwise comparisons of the luxury-necessity characteristics of the commodities in each of the three basic groups given in Table 1. This suggests that the errors associated with the prior estimates of goods within each group are positively correlated. We shall somewhat arbitrarily assign a value of 0.5 to the correlation of the estimates within each of the three blocks of goods.[15] All other error terms will be postulated to be uncorrelated since, for example, the prior estimate of the income elasticity of Groceries is essentially unrelated to that of, say, Clothing or Footwear.

If we define D as the 13 by 13 diagonal matrix whose diagonal elements consist of the first 13 standard deviations given in Table 1, column 4, the 13 by 13 covariance matrix (matrix of second moments around μ) can be written

$$V_0 = DAD \tag{10}$$

where A is the block-diagonal correlation matrix. Recalling that the prior estimate of σ_{14} was computed from the identity $\sum w_i\sigma_i = 1$, which implies $\hat{\mu}_{14} = 1 - \sum_{k=1}^{13} \hat{\mu}_k$, we can easily compute the variance and covariances of $\hat{\mu}_{14}$. The variance is $l'V_0l$ where l is a 13-element vector of units. The square root of this value, 0.347, is given in Table 1.

6. Parameterization and Estimation

Substituting the constraints (9.1) and (9.2) into (8), the demand model can be written

$$w_{it}^* Dq_{it} = \mu_i Dq_t + \varphi A_{it}(\mu) + \sum_{j \neq i} v_{ij}(Dp_{jt} - Dp_{it}) + \varepsilon_{it} \tag{11}$$

where $A_{it}(\mu) = \mu_i[(Dp_{it} - Dp_{nt}) - \sum_k^{13} \mu_k(Dp_{kt} - Dp_{nt})]$. Notice that all own-price coefficients are eliminated from (11) and that μ_{14} does not appear. It is easily shown (see e.g., Theil [25] pp. 54–56) that the n-th equation in the system (11) is an exact linear combination of the first $n-1$ equations. That equation ($i = 14$) may therefore be deleted and point estimates of the parameters of the deleted equation can be obtained from (9).[16] We then directly estimate the parameters of only the first 13 equations.

The price coefficient matrix $N = [v_{ij}]$ is of the form

$$N = \begin{bmatrix} N_1 & 0 \\ 0 & N_2 \end{bmatrix}$$

[15] Experiments were undertaken on the sensitivity of the mixed estimates and standard errors to changes in this correlation. It was found that both the point estimates and standard errors were quite robust under changes in the correlation.

[16] See Theil [24], Ch. 6.

where N_1 is an 11 by 11 diagonal matrix with ν_{ii}, $i = 1, \ldots, 11$ on the diagonal and N_2 is the 3 by 3 matrix

$$N_2 = \begin{bmatrix} \nu_{12,12} & \nu_{12,13} & \nu_{12,14} \\ \nu_{13,12} & \nu_{13,13} & \nu_{13,14} \\ \nu_{14,12} & \nu_{14,13} & \nu_{14,14} \end{bmatrix}$$

The block diagonal form of N follows from declaring the utility function to be strongly separable with each of the first 11 goods being preference independent and the remaining three goods forming a single preference dependent block (see Theil [25] and Paulus [15] for details). Imposing the symmetry restriction (9.3) on N we find that only three price coefficients remain—$\nu_{12,13}$, $\nu_{12,14}$ and $\nu_{13,14}$ (recall that all own-price coefficients are eliminated from (11) by the constraint (9.2)). The parameter vector to be estimated then consists of 17 elements which includes the first 13 marginal budget shares, the income flexibility, and the three off-diagonal price coefficients.

The model (11) is non-linear in the parameters with the source of non-linearity being the dependence of A_{it} (as defined below equation (11)) on the marginal budget shares. An iterative GLS procedure is therefore employed to obtain converged point estimates based on the sample evidence. See Theil [24], pp. 588–594 for details of the estimation procedure and pp. 596–598 for an asymptotic evaluation of the iterative GLS estimator.[17]

After converged sample estimates are obtained, mixed estimates can be easily computed from

$$\hat{\beta}_M = (X'\hat{\Omega}^{-1}X + R'V_0^{-1}R)^{-1}(X'\hat{\Omega}'y + R'V_0^{-1}r) \tag{12}$$

Where $R = [I_{13}\ 0]$, V_0 is defined in (10), $X = [X_t]$ is the 29×13 by 13 observation matrix on the independent variables (including converged values of μ_i in $A_{it}(\mu)$), and $\hat{\Omega} = I_{29} \otimes \hat{\Sigma}$, with $\hat{\Sigma}$ being the 13 by 13 matrix of mean squares and products of the converged GLS residuals.[18] The vector $y = [y_t]$ consists of 29×13 observations on the dependent variable and r is the 13-element vector of prior estimates of the marginal budget shares. Standard errors of the mixed estimates are computed by taking the square roots of the diagonal elements of the inverse on the right hand side of (12).

The sample and mixed estimates are given in Tables 2 and 3. The first two columns of Table 2 show the sample and mixed marginal shares. Notice that the gain in precision of the mixed estimates (as measured by the standard deviations) over those of the sample is sometimes considerable; see, for example, the estimates for Bread, Groceries, and Fish. Going back to column 4 of Table 1, we see that for all but two goods, Clothing and Other goods and services, the mixed estimates are between the prior and the sample estimates. It should also be noted that when

[17] Instead of using "prior" estimates obtained from the absolute price Rotterdam model as Theil [24] suggests, p. 591, we use the prior estimates given in Table 1 as starting values for the iterative estimator.

[18] $\Omega = I_{29} \otimes \Sigma$ implies serial independence in the disturbances. Tests of this hypothesis were carried out in Paulus [15] and the serial independence hypothesis was accepted.

TABLE 2

INCOME AND OWN-PRICE COEFFICIENTS

Commodity Group	Marginal share (μ_i) Sample (1)	Marginal share (μ_i) Mixed (2)	Income elasticity Prior (3)	Income elasticity Mixed (4)	Own-price coefficient (ν_{ii}) Sample (5)	Own-price coefficient (ν_{ii}) Mixed (6)
1. Bread	0.05 (0.68)	0.56 (0.27)	0.2 (0.10)	0.16 (0.08)	−0.03 (0.45)	−0.37 (0.18)
2. Groceries	5.67 (1.21)	2.80 (0.60)	0.4 (0.15)	0.50 (0.11)	−3.79 (0.82)	−1.38 (0.41)
3. Dairy products	3.27 (0.84)	3.51 (0.56)	0.6 (0.15)	0.49 (0.08)	−2.19 (0.57)	−2.35 (0.39)
4. Vegetables and fruit	3.33 (1.03)	2.76 (0.58)	0.6 (0.20)	0.62 (0.13)	−2.23 (0.70)	−1.85 (0.39)
5. Meat	5.17 (1.44)	6.76 (0.94)	1.0 (0.20)	0.89 (0.12)	−3.45 (0.97)	−4.53 (0.65)
6. Fish	0.49 (0.45)	0.50 (0.12)	0.8 (0.25)	0.71 (0.16)	−0.33 (0.30)	−0.34 (0.08)
7. Pastry, chocolate, ice cream	2.27 (0.68)	2.43 (0.41)	0.8 (0.25)	0.78 (0.13)	−1.52 (0.46)	−1.63 (0.28)
8. Beverages	3.16 (0.63)	3.82 (0.44)	1.5 (0.30)	1.41 (0.16)	−2.11 (0.42)	−2.56 (0.31)
9. Tobacco	3.01 (0.85)	2.34 (0.56)	0.6 (0.25)	0.62 (0.15)	−2.01 (0.57)	−1.57 (0.38)
10. Household and other durables	22.33 (1.50)	21.25 (1.06)	2.0 (0.50)	2.26 (0.11)	−14.92 (1.19)	−14.23 (0.94)
11. Water, light and heat	5.92 (1.11)	4.62 (0.76)	0.8 (0.30)	0.84 (0.14)	−3.95 (0.76)	−3.10 (0.53)
12. Clothing and other textiles	23.53 (2.03)	22.63 (1.59)	2.0 (0.40)	1.70 (0.12)	−9.43 (1.43)	−8.74 (1.17)
13. Footwear	2.97 (0.40)	2.38 (0.25)	1.5 (0.30)	1.66 (0.18)	−1.06 (0.31)	−0.65 (0.22)
14. Other goods and services	18.83 (2.77)	23.64 (2.00)	0.7 (0.35)	0.75 (0.06)	−5.52 (2.11)	−8.52 (1.69)

Note. The figures in columns (1), (2), (5) and (6) are all to be multiplied by 10^{-2}.

there is a sizable difference in the prior and the sample standard error of an estimate, the mixed estimate will usually be closer to the estimate with the smaller standard error; see the estimates for Bread, Groceries, Clothing, and Household and other durables.

In columns 3 and 4, the prior and mixed estimates of the income elasticities are given, where the mixed estimates and standard deviations are computed by dividing the mixed marginal shares and standard deviations by the corresponding average budget shares. The two columns show how the sample information modified our prior assessments of the income elasticities. What is particularly interesting, given that the estimates are quite close, is the difference in the size of the standard deviations. These differences range from about 25 per cent for Bread to a factor of from 3 to 5 for Clothing, Household and other durables, and Other goods and services, with the mixed estimates always being more precise.

The own-price coefficients are computed from the restriction (9.2), which implies

$$\nu_{ii} = \varphi\mu_i - \sum_{j \neq i}^{14} \nu_{ij}$$

For the preference independent goods, $i = 1, 2, \ldots 11$, the own-price coefficients all equal $\varphi\mu_i$, since the off-diagonal price coefficients are zero for these goods. The sample and mixed estimates and standard deviations are shown in columns 5 and 6.[19] The estimates of φ which are used to compute the own-price coefficients are

$$\hat{\varphi} = -0.6678(0.0272) \text{ sample}$$

$$\hat{\varphi} = -0.6696(0.0271) \text{ mixed.}$$

The sample and mixed estimates of the off-diagonal price coefficients are shown in Table 3. It is seen that these estimates are not affected very much by the introduction of stochastic prior information against the marginal shares. The estimates of $\nu_{12,14}$ and $\nu_{13,14}$ are noteworthy because they imply strong complementarity between Clothing and Other goods and services and Footwear and

TABLE 3

SAMPLE AND MIXED ESTIMATES OF OFF-DIAGONAL PRICE COEFFICIENTS

	Sample			Mixed		
	Clothing	Footwear	Other goods	Clothing	Footwear	Other goods
Clothing	−9.43	0.02	−6.31	−8.74	−0.02	−6.39
	(1.43)	(0.20)	(0.75)	(1.17)	(0.20)	(0.75)
Footwear		−1.06	−0.95		−0.65	−0.92
		(0.31)	(0.18)		(0.22)	(0.18)
Other goods and services			−5.52			−8.52
			(2.11)			(1.69)

Note. All figures are to be multiplied by 10^{-2}.

[19] See H. Theil [24], pp. 373–374 and 598–602 for details on asymptotic standard errors for functions of sample moments.

Other goods and services. This is not too surprising in view of the composition of Other goods and services which includes such things as taylor's services and shoe repair. The repair of clothing, carpets, (which is included in the Clothing composite good) and footwear was more common in the Netherlands during the sample period than it is today in the United States.

The two matrices shown in Table 3 are both negative definite. The sample and mixed 14 by 14 price matrices will then be negative definite, as required by cardinal utility theory,[20] since the own-price coefficients of the 11 preference independent goods are all negative.

7. Statistical Analysis of the Model

The gain in precision that was discussed in the last section was attributed to the use of stochastic prior information. In order to determine whether this information is compatible with the sample evidence, we compute the compatibility test statistic (7.1) given in Section 3. Since Ω is unknown, we use the approximation $\hat{\Omega} = I \otimes \hat{\Sigma}$, with $\hat{\Sigma}$ being the matrix of mean squares and products of the converged sample residuals, to compute (7.1). The statistic is then approximately distributed $\chi^2(13)$. The computed value of (7.1) is 13.72 which is not nearly significant for $\chi^2(13)$. We therefore accept the prior and sample information as being mutually compatible.[21]

After establishing the compatibility of the two sources of information, we shall determine the relative shares of the posterior precision contributed by each. To obtain values of α_p and α_s, we replace Ω by $\hat{\Omega}$ in (7.2) and use $k = 17$. We find

$$\alpha_p = 0.3719 \qquad \alpha_s = 0.6281$$

which indicates that the prior and sample information contribute about 37 and 63 percent, respectively, to the posterior precision of the point estimates.

It will now be shown that the product $\alpha_s k$ can be interpreted as the "effective number of unconstrained parameters" in the model after stochastic prior information is introduced. Suppose the model is

$$(13) \qquad y_t = X_t\beta + \varepsilon_t \qquad E\varepsilon_t = 0 \qquad V(\varepsilon_t) = \Sigma \qquad t = 1, 2, \ldots, T$$

and that prior estimates of some of the parameters are available. After the mixed estimator is computed, (13) becomes

$$(14) \qquad y_t = X_t\hat{\beta}_M + e_t$$

[20] Latent roots were computed for both matrices, and all were negative.

[21] Yancy, Bock, and Judge [29] evaluate the power of the compatibility statistic using Monte Carlo methods. They show that for 120 samples of 10, 15, and 25 observations, using a 5 percent significance level, the compatibility statistic falls in the rejection region 80, 118, and 120 times when biased prior estimates are used, where the parameter misspecification amounted to an error of one to two standard deviations on each of the six parameters estimated.

where $\hat{\beta}_M$ is the vector of mixed estimates and e_t is the corresponding vector of mixed residuals. Subtracting (13) from (14) then gives

$$e_t = \varepsilon_t - X_t(\hat{\beta}_M - \beta) \tag{15}$$

Post-multiplying (15) by e_t', summing over t, dividing by T and taking the expectation, we find the expectation of the matrix of mean squares and products of the mixed residuals

$$E\left(\frac{1}{T}\sum_t^T e_t e_t'\right) = \Sigma - \frac{1}{T}\sum_t^T X_t(X'\Omega^{-1}X + R'V_0^{-1}R)^{-1}X_t' \tag{16}$$

where $E(\hat{\beta}_M - \beta)(\hat{\beta}_M - \beta)' = (X'\Omega^{-1}X + R'V_0^{-1}R)^{-1}$ is used. Equation (16) shows that the expectation of $1/T\sum_t^T e_t e_t'$ is exceeded by the contemporaneous covariance matrix Σ by a positive semidefinite matrix.

Since it is impossible to find a multiplicative correction factor which insures that each element of $1/T\sum e_t e_t'$ is an unbiased estimate of the corresponding element of Σ, we confine ourselves to an unbiasedness correction for a scalar function of this matrix. Suppose we post-multiply $E\hat{\Sigma}$ by Σ^{-1}, where $\hat{\Sigma}$ is now the matrix of mean squares and products of the mixed residuals apart from the unknown multiplicative correction factor. If $E\hat{\Sigma} = \Sigma$ the trace of this product will be 13. We therefore impose $tr\, E(\hat{\Sigma}\Sigma^{-1}) = 13$ on $\hat{\Sigma}$. To derive the correction factor it is sufficient to post-multiply (16) by Σ^{-1} and take the trace of both sides,

$$\begin{aligned}
E\,\mathrm{tr}\frac{1}{T}\sum_t^T e_t e_t'\Sigma^{-1} &= \mathrm{tr}\,\Sigma\Sigma^{-1} \\
&\quad - \mathrm{tr}\left(\frac{1}{T}\sum_t^T X_t(X'\Omega^{-1}X + R'V_0^{-1}R)^{-1}X_t'\Sigma^{-1}\right) \\
&= \mathrm{tr}\, I_{13} - \frac{1}{T}(X'\Omega^{-1}X + R'V_0^{-1}R)^{-1}\left(\sum_t X_t'\Sigma^{-1}X_t\right) \\
&= 13 - \frac{1}{T}tr\,(X'\Omega^{-1}X)(X'\Omega^{-1}X + R'V_0^{-1}R)^{-1} \\
&= 13 - \frac{k}{T}\alpha_s
\end{aligned}$$

where tr AB = tr BA is used after the second and third equal signs. The equality of $(X'\Omega^{-1}X)$ and $\sum_t^T X_t'\Sigma^{-1}X_t$ follows from $\Omega = I \otimes \Sigma$ and $X = [X_t]$.[22]

To obtain $E(\mathrm{tr}\,1/T\sum_t e_t e_t'\Sigma^{-1}) = 13$, we multiply $1/T\sum_t e_t e_t'$ by $13/(13 - \alpha_s k/T)$. The corrected estimate of the covariance matrix of the disturbances is then

$$\hat{\Sigma} = \frac{13T}{13T - \alpha_s k}\frac{1}{T}\sum_t^T e_t e_t' \tag{17}$$

This leads to the following interpretation of the product $\alpha_s k$. If no stochastic prior information is introduced against the parameters, the analogous correction will be $13T/(13T - k)$. Comparing this with equation (17), it is seen that $\alpha_s k$ plays the

[22] Recall that, for this study, the hypothesis of serial independence of the demand disturbances was acceptable.

role of the number of unconstrained parameters in the correction for loss of degrees of freedom after stochastic prior information is introduced. It is therefore natural to regard $\alpha_s k$ as the "effective number of unconstrained parameters" in the model after probabilistic constraints are imposed.

For the 17 parameter model, with $\alpha_s = 0.6281$. we have $\alpha_s k = 10.7$. The use of stochastic prior information on the marginal budget shares has then led to a reduction in the effective number of unconstrained parameters from 17 to 10.7.

8. Summary

In this study we have shown how stochastic prior information on a subset of the parameter vector can be effectively combined with the sample evidence to yield posterior estimates that are much more precise than those of the sample. The method of mixed estimation is used to combine the prior and sample information because of its flexibility and ease of implementation. It is determined that the prior information is compatible with the evidence of the sample and it appears that the share of the posterior precision that can be attributed to the prior information is a little over 35 percent. It was shown that this amounts to a reduction in the effective number of unconstrained parameters in the model from 17 to 10.7.

Board of Governors
Federal Reserve System

References

[1] Barten, A. P.. "Consumer Demand Functions Under Conditions of Almost Additive Preferences." *Econometrica*. Vol. 32. 1964.
[2] Barten. A. P.. "het verbruik door gezinshuishoudingen in Nederland. 1921–39 en 1948–62." Report 6604 of the Econometric Institute of the Netherlands School of Economics. Rotterdam. 1966.
[3] Barten. A. P.. "Estimating Demand Equations." *Econometrica*. Vol. 36. 1968.
[4] Barten. A. P.. "Maximum Likelihood Estimation of a Complete System of Demand Equations." *European Economic Review*. Vol. 1. 1969.
[5] Byron. R. P.. "The Restricted Aitken Estimation of Sets of Demand Relations." *Econometrica*. Vol. 38. 1970.
[6] Chipman. J. S.. "On Least Squares with Insufficient Observations." *Journal of the American Statistical Association*. Vol. 59. 1964.
[7] Ericson. W. A.. "A Note on the Posterior Mean of a Population Mean." *Journal of the Royal Statistical Society*. Series B. Vol. 31. 1969.
[8] Ferguson. T. S.. *Mathematical Statistics: A Decision Theoretic Approach*. New York: Academic Press. 1967.
[9] Frazer. D. A. S.. "The Fiducial Method and Invariance." *Biometrika*. 1961.
[10] Frazer. D. A. S.. "On Fidicial Inference." *Annals of Mathematical Statistics*. Vol. 32. 1961.
[11] Kaufman. G. M.. "Comments." pp. 204–207. in M. D. Intriligator (ed.) *Frontiers of Quantitative Economics*. Amsterdam: North-Holland Publishing Co.. 1971.
[12] Mehta. J. S. and P. A. V. B. Swamy. "The Finite Sample Distribution of Theil's Mixed Regression Estimator and a Related Problem." *Review of the International Statistical Institute*. Vol. 38. 1970.
[13] Mehta. J. S. and P. A. V. B. Swamy. "The Exact Finite Sample Distribution of Theil's Compatibility Statistic and its Application." *Journal of the American Statistical Association* (forthcoming).
[14] Mehta. J. S. and P. A. V. B. Swamy. "Efficient Method of Estimating the Level of a Stationary First-Order Autoregressive Process." *Communications in Statistics* (forthcoming).
[15] Paulus. J. D.. "The Estimation of Large Systems of Consumer Demand Equations Using Stochastic Prior Information." Doctoral Dissertation. University of Chicago. 1972.

[16] Prais. S. J. and H. S. Houthakker. *An Analysis of Family Budgets*. Cambridge. 1955.
[17] Rao. C. R.. *Linear Statistical Inference and Its Applications*. New York: John Wiley and Sons. 1965.
[18] Rothenberg. T. J.. "A Bayesian Analysis of Simultaneous Equation Systems." Report 6315. Econometric Institute of the Netherlands School of Economics. Rotterdam. 1963.
[19] Swamy. P. A. V. B. and J. S. Mehta. "On Theil's Mixed Regression Estimator." *Journal of the American Statistical Association*. Vol. 64, 1969.
[20] Swamy. P. A. V. B. and J. S. Mehta. "Relative Efficiencies of Several Biased and Unbiased Estimators of Regression Coefficients." (mimeographed) 1973.
[21] Theil. H.. "On the Use of Incomplete Prior Information in Regression Analysis." *Journal of the American Statistical Association*. Vol. 58. 1963.
[22] Theil. H.. "The Information Approach to Demand Analysis." *Econometrica*. Vol. 33. 1965.
[23] Theil. H.. *Economics and Information Theory*. Amsterdam: North-Holland. 1967.
[24] Theil. H.. *Principles of Econometrics*. New York: John Wiley and Sons. and Amsterdam: North-Holland. 1971.
[25] Theil. H.. "Introduction to Demand and Index Number Theory." Report 7204. Center for Mathematical Studies in Business and Economics. University of Chicago. 1972.
[26] Theil. H.. "Mixed Estimation Based on Quasi-Prior Judgements." Report 7306. Center for Mathematical Studies in Business and Economics. University of Chicago. 1973.
[27] Theil. H. and A. S. Goldberger. "On Pure and Mixed Statistical Estimation in Economics." *International Economic Review*. Vol. 2. 1961.
[28] Wold. H.. *Demand Analysis*. New York: John Wiley and Sons and Stockholm: Almgvist and Wiksell. 1953.
[29] Yancy. T. A.. Bock. M. E.. and G. G. Judge. "Finite Sample Results from Theil's Mixed Regression Estimator." *Journal of the American Statistical Association*. Vol. 67. 1972.
[30] Zellner. A.. *An Introduction to Bayesian Inference in Econometrics*. New York: John Wiley and Sons. 1971.