5

Stochastic Specifications

In this chapter we outline several stochastic specifications that are important in empirical demand analysis and that form the basis of our empirical results in Chapters 6 and 7. In Section 1 we develop what we refer to as our standard stochastic specification for demand systems written in share form: additive, independent (across observations but not goods), normal errors with 0 mean and a constant nondiagonal contemporaneous covariance matrix. In Section 2 we extend this standard model to allow for first order vector autoregressive systems. We pay particular attention to the effects of including the initial observation in the estimation procedure. We consider two procedures that differ in their treatment of the initial observation: the generalized first difference method in which the initial observation is used only in differencing, and an alternative method in which full use is made of the initial observation. Almost all consumer demand studies in the literature that incorporate serial correlation do so using the generalized first difference method. In Section 3 we develop an error components model in the context of a time series of cross sections. This extends our standard specification to one in which disturbances have a "time-specific" component in addition to a general component. We demonstrate that the resulting likelihood function can be written in a computationally tractable form even though it cannot in general be concentrated with respect to any of the disturbance covariance parameters. Finally in Section 4 we investigate the possibility of estimating random coefficients models for the LES, quadratic expenditure system (QES), and basic translog (BTL) models. We show that estimation of such models is computationally tractable for the LES and QES if the appropriate subset of parameters is assumed to be random. However, the translog forms are less tractable, particularly with more than two goods, and their estimation is beyond the scope of this book.

1. THE STANDARD STOCHASTIC SPECIFICATION

We begin with the simplest, and what we will refer to as our "standard," stochastic specification. We write our system of demand equations for observation t in share form as

$$\text{(1)} \qquad w_{it} = \omega^i(z_t, \beta) + u_{it}, \qquad \begin{array}{l} i = 1, \dots n \\ t = 1, \dots T, \end{array}$$

where w_{it} is the share of the ith good in total expenditure, z_t is the set of explanatory variables, β the parameters to be estimated, and u_{it} a random disturbance. Denoting the row vector $(u_{1t}, \dots, u_{nt})$ as $\tilde{u}_t'$ we assume that $E(\tilde{u}_t\tilde{u}_t') = \tilde{\Omega}$ and that $E(\tilde{u}_t\tilde{u}_s) = 0$ for $s \neq t$. That is, we assume that the contemporaneous covariance matrix for the share disturbances is the same for all observations, and that the disturbances are uncorrelated across observations. We relax the latter assumption in Section 2 when we introduce autocorrelation in the context of a time series model.

The constant covariance assumption appears plausible a priori in view of the fact that the dependent variables are shares, and thus bounded by 0 and 1. On the other hand, if one adds disturbance terms to the demand equations in expenditure form, the assumption of a constant covariance matrix is less plausible. In a cross-section context it implies that the variance of the disturbance associated with expenditure on a good will be the same regardless of the level of expenditure, which may vary widely within a sample. In a time series context it implies that increases in per capita consumption over time will not be accompanied by increases in the disturbances' variances. Furthermore this specification implies that if all prices and total expenditure were to increase proportionately, then the variances of the expenditure equation disturbances would remain constant and thus the variances of the disturbances of the demand equations in share form would decrease, even though the predicted shares remained constant. On theoretical grounds we prefer a specification in which the covariance matrix of the disturbances associated with the demand equations in share form is unaffected by proportional changes in all prices and expenditure. An alternative approach is to assume that the covariance matrix of the disturbances associated with the expenditure equations is proportional to the square of expenditure. This is the approach followed, for example, in Wales [1971]. Finally if one adds disturbance terms to the demand equations in quantity form, then the assumption of a constant covariance matrix for the disturbances is entirely inappropriate. Indeed if there are n linearly independent prices in the sample then this assumption implies that the covariance matrix is of rank 0. For a proof of this proposition see, for example, Pollak and Wales [1969].

The fact that expenditure on the n goods exhausts the budget imposes a restriction on $\tilde{\Omega}$. In particular, summing (1) over all goods gives

$$\text{(2)} \qquad \sum w_{kt} = \sum \omega^k(z_t, \beta) = 1,$$

implying that $\sum_k u_{kt} = 0$ for each t, and thus the u_{kt}'s cannot be mutually independent. Further for each t we have $E(\tilde{u}_t\tilde{u}_t') = \tilde{\Omega}$ and $\ell'\tilde{u}_t = 0$, where $\ell' = (1, \dots, 1)$ a vector of n ℓ's. This implies that $\ell'\tilde{\Omega} = 0$, in which case ℓ' is an eigenvector of $\tilde{\Omega}$ with the corresponding eigenvalue equal to 0; thus $\tilde{\Omega}$ is singular.

We assume that $\tilde{u}_t$ has a multivariate normal distribution with mean 0 and covariance matrix $\tilde{\Omega}$ for all t. Due to the singularity of $\tilde{\Omega}$ the density for $\tilde{u}_t$ may be expressed in terms of the density of any $n-1$ of the goods. Barten [1969] proves that the parameter estimates are independent of the choice of good deleted. This is a major advantage of the maximum likelihood procedure over a two-step Zellner-type procedure for which the estimates depend on the choice of good deleted. We arbitrarily drop the nth good and define $u_t' = (u_{1t}, \ldots, u_{n-1,t})$ and the corresponding covariance matrix as $E(u_t u_t') = \Omega$. Under these assumptions the density for u_t is given by

$$f(u_t) = (2\pi)^{-[(n-1)/2]}|\Omega|^{-1/2}\exp\left(-\frac{u_t'\Omega^{-1}u_t}{2}\right), \tag{3}$$

and the logarithm of the likelihood function for a sample of T (independent) observations is given by

$$L(\beta, \Omega) = -\frac{(n-1)T}{2}\log 2\pi - \frac{T}{2}\log|\Omega| - \frac{1}{2}\sum_{t=1}^{T} u_t'\Omega^{-1}u_t. \tag{4}$$

For estimation purposes it is convenient to concentrate this likelihood function with respect to the elements of Ω. Following, for example, Rothenberg and Leenders [1964, p. 61], the concentrated logarithm of the likelihood may be expressed as

$$L(\beta) = -\frac{(n-1)T}{2}(\log 2\pi + 1) - \frac{T}{2}\log|S|, \tag{5}$$

where S is a square matrix of order $n-1$, with the ijth element given by

$$s_{ij} = \left(\sum_{t=1}^{T} u_{it}u_{jt}\right)\Big/ T, \qquad i,j = 1, \ldots, n-1. \tag{6}$$

Thus S is just the sample covariance matrix of the residuals for the first $n-1$ goods. Maximizing the likelihood function is equivalent to minimizing the determinant of S, which is a function only of the β's and the data. This minimization may be carried out using standard nonlinear minimization algorithms. All of the empirical results we present in Chapters 6 and 7 are based on maximizing various likelihood functions using an algorithm due to Fletcher [1972] and rely on numerically calculated derivatives. The material in the first three sections of Chapter 6 is based on maximizing a likelihood function of the form given by (5).

2. MODELS WITH FIRST ORDER AUTOCORRELATED ERRORS

In this section we extend the standard stochastic specification of Section 1 by dropping the assumption of independence across observations. We assume that the $\tilde{u}_t$ follow a first order autoregressive process

$$\tilde{u}_t = \tilde{R}\tilde{u}_{t-1} + \tilde{e}_t, \qquad t = 2, \ldots, T, \tag{7}$$

where the $\tilde{e}_t$ are independent and normally distributed with mean 0 and a constant contemporaneous covariance matrix. Once again the presence of the budget constraint implies restrictions on some of the covariance parameters. It can be shown that only $(n-1)^2$ independent transformations of the $\tilde{R}$ matrix can be identified. Denote the transformed $\tilde{R}$ matrix by R. Then R is obtained from $\tilde{R}$ by first subtracting the last column of $\tilde{R}$ from each of the other columns of $\tilde{R}$, and then deleting the last row and column. Berndt and Savin [1975] discuss this transformation procedure. Thus we have the following first order autoregressive process

$$u_t = Ru_{t-1} + e_t, \qquad t = 2, \ldots, T, \tag{8}$$

where u_t and e_t are both $(n-1) \times 1$ vectors, and the e_t are independently normally distributed with 0 mean and constant covariance matrix Ω. We assume that the process is stationary, implying the u_t are normally distributed with mean 0 and contemporaneous covariance matrix θ, which from (8) satisfies

$$\theta = R\theta R' + \Omega. \tag{9}$$

We consider first a procedure that we refer to as the "generalized first difference" procedure, a straightforward generalization to a system of equations of the standard first difference procedure used in the single equation context. This generalized first difference procedure premultiplies the system of $n-1$ equations by R, lags once, and subtracts from the original observation. This yields a system of equations with independent disturbances with mean 0 and covariance matrix Ω. Maximum likelihood estimates of this differenced system do not coincide with maximum likelihood estimates based on the original system and sample: the generalized first difference procedure fails to make full use of the first observation, using it only in differencing. However, the generalized first difference procedure offers a major computational advantage: estimation is relatively straightforward because the likelihood function corresponding to the differenced system can be concentrated and maximized in the usual way. On the other hand, maximum likelihood estimates corresponding to the original system and sample are difficult to calculate because there appears to be no way of concentrating the likelihood function.[1]

With this first differencing procedure the log likelihood has the same form as (4) with $u_t - Ru_{t-1}$ replacing u_t, and $T-1$ replacing T. Further since $\Omega = \theta - R\theta R'$ from (9) we can write the log likelihood as a function of θ, β, and R as follows:

[1] Beach and MacKinnon [1979] discuss the problem of concentrating the likelihood function.

$$(10)\quad L(\theta,\beta,R)=k_1-\left(\frac{T-1}{2}\right)\log|\theta-R\theta R'|-\frac{1}{2}\sum_{t=2}^{T}(u_t-Ru_{t-1})'$$
$$\cdot(\theta-R\theta R')^{-1}(u_t-Ru_{t-1})$$

which, as in the case of (4), can be concentrated with respect to the covariance parameters θ to give

$$(11)\qquad L(\beta,R)=k_2-\left(\frac{T-1}{2}\right)\log\left|\sum_{t=2}^{T}(u_t-Ru_{t-1})(u_t-Ru_{t-1})'\right|,$$

where k_1 and k_2 are constants independent of θ, β, and R.[2]

Maximization of the likelihood given by (11) is straightforward but slightly more complicated than that given by (4), because the likelihood in (11) depends on the additional parameters contained in the R matrix.

A special case of (11) involves a diagonal R matrix. When R is diagonal, the random disturbance associated with any good depends only on the lagged value of the disturbance for that good, but is independent of lagged disturbances associated with other goods. For each good we replace (8) by

$$(12)\qquad u_{it}=\rho_i u_{it-1}+e_{it},\qquad i=1,\ldots,n;\qquad t=2,\ldots,T,$$

where ρ_i is the serial correlation coefficient associated with the ith good. In this case, however, it can be shown that in order to ensure that the estimates are independent of the equation deleted, all the ρ_i must be equal (because of restrictions resulting from the budget constraint). Berndt and Savin [1975] give a proof of this proposition. Thus, introducing a diagonal R matrix results in only one additional parameter to be estimated. This is attractive from a computational point of view but may be too restrictive to model adequately the true autocorrelation process.

We consider next maximum likelihood estimation of the entire system, a procedure that takes full account of the first observation instead of using it merely to calculate first differences. We assume that $u_1=Se_1$ where S is defined to ensure the stationary of the u process, that is,

$$(13)\qquad \theta=S\Omega S'.$$

The log likelihood function for all T observations is then

$$(14)\quad L(\theta,\beta,R)=k_3-\frac{T}{2}\log|\theta-R\theta R'|+\tfrac{1}{2}\log(|\theta-R\theta R'|/|\theta|)$$
$$-\frac{1}{2}\left[u_1'\theta^{-1}u_1+\sum_{t=2}^{T}(u_t-Ru_{t-1})'(\theta-R\theta R')^{-1}(u_t-Ru_{t-1})\right]$$

[2]The matrix whose determinant is taken in (11) differs by the factor $T-1$ from that in (4), thus giving rise to offsetting differences in the constant terms. We have done this to make our presentation consistent with that of Beach and MacKinnon [1979] on which our analysis is based.

where k_3 is a constant. As noted by Beach and MacKinnon, (14) differs from (10) in two respects: it contains the additional term $-(u_1'\theta^{-1}u_1)/2$ and an additional term that constrains $|\theta - R\theta R'|$ and $|\theta|$ to have the same sign $[\frac{1}{2}\log(|\theta - R\theta R'|/|\theta|)]$. The latter condition is equivalent to requiring the error process (8) to be stationary. That is, including the first observation explicitly constrains the elements of R to guarantee stationarity.

The log likelihood function (14) can be concentrated with respect to θ only when R is a diagonal matrix with a common element ρ on the diagonal. But as mentioned above, when R is diagonal and a system of share or expenditure equations is being estimated, then a diagonal R matrix is restricted to one in which all diagonal elements are the same. In this case (14) becomes

$$(15) \qquad L(\beta,\rho) = k_4 + \frac{(n-1)}{2}\log(1-\rho^2) - \frac{T}{2}\log|(1-\rho^2)u_1u_1' + \sum_{2}^{T}(u_t - \rho u_{t-1})(u_t - \rho u_{t-1})'|$$

where k_4 is a constant.

Maximizing (15) is not much more difficult computationally than maximizing the log likelihood given by (4). However, the unconcentrated log likelihood function given by (14) involves an additional $n(n-1)/2$ unknown parameters as compared with the log likelihood given by (4), and this may represent a significant additional computational burden, especially when the number of goods in the system is large and the share equations are highly nonlinear.

Finally maximization of (14) requires that θ be positive definite. This restriction may be imposed by writing θ as the product of a lower triangular matrix θL and its transpose, and estimating the elements of θL rather than θ.[3]

The empirical results in Chapter 7 are based on maximizing likelihood functions of the form given by (11), (14), and (15).

3. A MODEL WITH TWO ERROR COMPONENTS

When working with a time series of cross sections, a natural extension of the standard stochastic specification is one in which the errors have more than one component. We consider here the case of two error components. Let u_{rt} denote the $(n-1) \times 1$ vector of disturbances added to $n-1$ share equations in a demand system, where, for example, r indexes a household

[3] Diewert and Wales [1987] discuss this technique in the context of imposing concavity on cost functions.

in a cross section at time t. These disturbances are assumed to be the sum of two components

$$u_{rt} = e_t + \varepsilon_{rt}, \qquad \begin{array}{l} r = 1, \ldots, q_t \\ t = 1, \ldots, T, \end{array} \tag{16}$$

where q_t is the number of households in the period t cross section and T is the number of cross sections. Note that if $r = 1$ for all t then this model reduces to the one discussed in Section 1 and contains T observations. But if $r > 1$ in one or more time periods then the total number of observations will be $\sum_{t=1}^{T} q_t$, which exceeds T. The subscript r does not necessarily denote the same household in successive time periods, nor are there necessarily the same number of households in every period. The $(n-1) \times 1$ disturbance vectors e_t and ε_{rt} are assumed to be independently normally distributed with 0 means and covariances given by Γ and Δ, respectively, where Γ is positive semidefinite and Δ is positive definite. The covariance matrix for the u's is thus

$$E(u_{rt}u_{s\tau}) = \begin{cases} \Gamma + \Delta & r = s, \quad t = \tau \\ \Gamma & r \neq s, \quad t = \tau \\ 0 & \qquad\quad t \neq \tau. \end{cases} \tag{17}$$

In order to write out the likelihood for this model we define

$$V_t = (u_{1t}, \ldots, u_{q_t t}), \tag{18}$$

which is an $(n-1) \times q_t$ matrix, and

$$v_t = \operatorname{vec} V_t, \tag{19}$$

which is an $(n-1)q_t \times 1$ vector. Then using (17) we have

$$\Omega_t \equiv E(v_t v_\tau') = \begin{cases} S_t S_t' \otimes \Gamma + I_t \otimes \Delta, & t = \tau \\ 0 & t \neq \tau, \end{cases} \tag{20}$$

where $S_t' = (1, \ldots, 1)$ is a $1 \times q_t$ vector of 1's, and I_t is the identity matrix of order q_t. This error components model is now in essentially standard form, and thus the log likelihood function for a sample of T independent time periods is given by (aside from an additive constant)

$$L(\beta, \Gamma, \Delta) = -\frac{1}{2}\left\{ \sum_{t=1}^{T} \log|\Omega_t| + \sum v_t'\Omega_t^{-1}v_t \right\} \tag{21}$$

with Ω given by (20) and v_t by (19). This likelihood function can now be maximized with respect to the elements of β, Γ, Δ. For estimation purposes it would be convenient if this expression could be concentrated with respect to the Γ and Δ parameters. Although this is possible for the case when the q_t's are equal for all t, it does not appear to be possible when the q_t's differ.

Since Γ and Δ are $(n-1)\times(n-1)$ symmetric matrices our stochastic specification contains $n(n-1)$ independent parameters in addition to the parameter set β. For large n and q_t, (21) may involve a prohibitive number of parameters to be estimated in view of the fact that Ω_t is of order $(n-1)q_t$. Fortunately (21) can be written in a simple form involving matrices of order $n-1$. We may rewrite Ω_t as

$$\Omega_t = \frac{S_tS_t'}{q_t} \otimes W_t + \left(I - \frac{S_tS_t'}{q_t}\right) \otimes \Delta, \tag{22}$$

where $W_t = \Delta + q_t\Gamma$, in which case it can be shown that[4]

$$\Omega_t^{-1} = \frac{S_tS_t'}{q_t} \otimes W_t^{-1} + \left(I - \frac{S_tS_t'}{q_t}\right) \otimes \Delta^{-1} \tag{23}$$

and

$$|\Omega_t| = |W_t|\,|\Delta|^{q_t-1}. \tag{24}$$

Further, from the properties of the trace operator and making use of (23) we may write $v_t'\Omega_t^{-1}v_t$ as

$$\begin{aligned} v_t'\Omega_t^{-1}v_t &= \operatorname{tr}\left[V_t\frac{S_tS_t'}{q_t}V_t'\right]W_t^{-1} + \operatorname{tr}\left[V_t\left(I - \frac{S_tS_t'}{q_t}\right)V_t'\right]\Delta^{-1} \\ &= \operatorname{tr}\left[\frac{V_tS_tS_t'V_t'(W_t^{-1}-\Delta^{-1})}{q_t}\right] + \operatorname{tr}(V_tV_t'\Delta^{-1}) \\ &= \frac{S_t'V_t'(W_t^{-1}-\Delta_t^{-1})V_tS_t}{q_t} + \operatorname{tr}(V_t'\Delta^{-1}V_t) \\ &= \sum_{s=1}^{q_t}\sum_{r=1}^{q_t} u_{st}'\frac{(W_t^{-1}-\Delta^{-1})}{q_t}u_{rt} + \sum_{s=1}^{q_t} u_{st}'\Delta^{-1}u_{st}. \end{aligned} \tag{25}$$

Substituting (25) and (23) into (21) yields the following log likelihood function (aside from an additive constant):

$$\begin{aligned} L(\beta,\Gamma,\Delta) = -\frac{1}{2}\sum_{t=1}^{T}\Bigg\{&\log|W_t| + (q_t-1)\log|\Delta| + \sum_{s=1}^{q_t} u_{st}'\Delta^{-1}u_{st} \\ &+ \sum_{s=1}^{q_t}\sum_{r=1}^{q_t} u_{st}'\frac{(W_t^{-1}-\Delta^{-1})}{q_t}u_{rt}\Bigg\}. \end{aligned} \tag{26}$$

The computational advantage of (26) lies in the fact that it contains matrices of order $(n-1)$, whereas (21) contains matrices of order $(n-1)q_t$. Thus even with large cross sections it is feasible to maximize the log likelihood function given by (26). However, since (26) cannot in general be concentrated, the $n(n-1)$ covariance parameters must be estimated

[4] A proof appears in Lemma 2.1 in Magnus [1982].

along with the demand system parameters. On the other hand, if $q_t = q$ for all t, that is, if each cross section contains the same number of observations, then the likelihood function can be concentrated with respect to Γ and Δ. In this case maximizing the likelihood is no more difficult computationally than maximizing the likelihood in the standard model of Section 1. For a proof see Magnus [1982, p. 251].

We have focused here on a model with two error components because our empirical work in Section 4 of Chapter 6 involves estimating such a model. If the cross section time series data are in panel form, then a model with three error components could be estimated. These components would reflect time-specific and household-specific effects in addition to the standard additive disturbance. Such a model appears, for example, in Avery [1977] and in Baltagi [1980].

4. RANDOM COEFFICIENTS MODELS

In this section we explore the possibility of incorporating a random coefficients stochastic structure into demand systems. In theory any subset of parameters can be stochastic; in practice, unless the stochastic subset of parameters is carefully chosen, the resulting demand system will be computationally intractable. Furthermore, each system must be studied separately to determine its tractability. We consider first the QES and LES models, and then the translog models.

4.1. Quadratic Expenditure System

The nonstochastic Σ-QES demand equations in share form (after deleting the observation subscript) are

$$(27)\qquad w_i = \frac{p_i b_i}{\mu} + a_i\left(1 - \frac{\sum p_k b_k}{\mu}\right) + \left(\frac{p_i}{\mu}c_i - a_i \sum \frac{p_k}{\mu} c_k\right) \prod \left(\frac{p_k}{\mu}\right)^{-2a_k} \left(1 - \sum \frac{p_k}{\mu} b_k\right)^2.$$

Suppose we now consider replacing a set of parameters θ_i (such as the a_i's, b_i's or c_i's) by $\theta_i + \varepsilon_i$, where ε_i is a normal random variable with mean 0. If this is done for the a's or b's in (27) the resulting error term is computationally intractable and, in particular, it is no longer normal. However, if we replace c_i in (27) by $c_i + \varepsilon_i$, then the resulting disturbance in (27), defined as v_i, is[5]

[5]Eq. (27) can be thought of as having been derived by maximizing the corresponding stochastic QES, where the ε_i's are known by the individual, but not by the researcher. The estimated β's are then the mean β's in the population. However, this is neither necessary nor especially useful. We believe that the specification of the error structure should be judged on its implications for the stochastic demand functions.

$$(28) \qquad v_i = \left(\frac{p_i}{\mu}\varepsilon_i - a_i \sum \frac{p_k}{\mu}\varepsilon_k\right)\delta(P,\mu), \qquad i = 1,\ldots,n$$

where

$$\delta(P,\mu) \equiv \prod \left(\frac{p_k}{\mu}\right)^{-2a_k} \left(1 - \sum \frac{p_k b_k}{\mu}\right)^2.$$

Since the v_i's are linear transformations of the ε_i's they are also normally distributed. The v_i's have 0 mean and covariance matrix Ω with ijth element

$$(29) \qquad \Omega_{ij} \equiv E(v_i v_j) = \frac{\delta^2(P,\mu)}{\mu^2}\Big[p_i p_j E(\varepsilon_i \varepsilon_j) - p_i a_j \sum p_k E(\varepsilon_i \varepsilon_k) - a_i p_j \sum p_k E(\varepsilon_k \varepsilon_j) + a_i a_j \sum \sum_{\ell} p_k p_\ell E(\varepsilon_k \varepsilon_\ell)\Big].$$

Further $\sum v_k = 0$ and we have the errors summing to 0 automatically.

Assumptions about $E(\varepsilon_i \varepsilon_j)$ will be based on their implications for Ω, the covariance matrix associated with the share equation disturbances. First we require that Ω be homogeneous of degree 0 in prices and expenditure. If all prices and expenditure increase proportionately, then the nonstochastic shares remain constant and it seems reasonable to assume the error variances do also. Second we assume that the share equation disturbances are homoskedastic with respect to expenditure. Although the homoskedasticity assumption is not necessary, it seems reasonable a priori, and implies an error structure that closely parallels that studied in Section 1 in which the error covariance matrix is independent of expenditure and prices. In fact we shortly relax this homoskedasticity assumption.

Perhaps the simplest way to obtain a covariance matrix for the share equation disturbances with these properties is to assume that the ε's satisfy

$$(30) \qquad E(\varepsilon_i \varepsilon_j) = \frac{\sigma_{ij}\mu^2}{p_i p_j \delta^2(P,\mu)}, \qquad i,j = 1,\ldots,n,$$

where the σ_{ij} are constant over observations. The resulting ijth element of Ω is then given by

$$(31) \qquad \Omega_{ij} \equiv E(v_i v_j) = \frac{p_i p_j \sigma_{ij} - p_i a_j \sum p_k \sigma_{ki} - a_i p_j \sum p_k \sigma_{kj} + a_i a_j \sum \sum_{\ell} p_k p_\ell \sigma_{k\ell}}{p_i p_j}.$$

Thus Ω is homogeneous of degree 0 in prices and expenditure together and, because it is independent of expenditure, it is also homogeneous of degree 0 in prices alone. These properties are attractive from a theoretical point of view, because the dependent variables are shares. From a computational point of view this error structure is less tractable than the standard formulation of Section 1 because the elements of Ω depend on the sample observations through the prices; hence the likelihood function cannot be concentrated with respect to the σ's.

Because the v's sum to 0 automatically, the budget constraint imposes no additional restrictions on the ε's. Thus, this is a richer specification than the standard stochastic specification in the sense that there are $n(n+1)/2$ independent covariance parameters (σ's), whereas the standard specification contains $n(n-1)/2$, yielding a difference of n between the two models. A special case of the random coefficients model is one in which $\sigma_{ij}=0$ for $i \neq j$, resulting in n independent variances. With $n=3$ this model contains the same number of covariances as does the standard stochastic specification, and for $n>3$ it contains fewer. Because the likelihood function cannot be concentrated in the random coefficients model, this restriction may be attractive on computational grounds. Of course in theory the random coefficients and the standard models may be used simultaneously. In this case if the c_i in (27) are replaced by $c_i+\varepsilon_i$, and an additional error term u_i, independent of ε_i, is appended to (27), then the ijth element of the disturbance covariance matrix for the system will be given by (31) plus a constant term (corresponding to the covariances of the u's). This covariance matrix contains $n(n+1)/2+n(n-1)/2=n^2$ independent elements, and in empirical applications it may well be the case that the data cannot distinguish this many distinct covariance parameters.

We can easily relax the assumption that the share equation disturbances are homoskedastic with respect to expenditure while maintaining the assumption of homogeneity of degree 0 in expenditure and prices. A particularly simple method and one that introduces only one additional parameter (α) is to replace (30) by

$$\text{(32)} \qquad E(\varepsilon_i\varepsilon_j)=\frac{\sigma_{ij}\mu^{2\alpha}}{p_i^{\alpha}p_j^{\alpha}\delta^2(P,\mu)}, \qquad i,j=1,\ldots,n.$$

Two special cases of (32) are of interest. If $\alpha=1$ Eq. (32) reduces to (30) and we have the model just considered, in which share equation disturbances are homoskedastic with respect to expenditure. If $\alpha=0$ then instead of (31) we have for the share equation disturbances, denoted by v_i^*

$$\text{(33)} \qquad E(v_i^* v_j^*)=\frac{p_i p_j}{\mu^2}\Omega_{ij},$$

where Ω_{ij} is given by (31). Since v_i^* is a share equation disturbance, (32) implies that the expenditure equation disturbances ($v_i^*\mu$) are homoskedastic with respect to expenditure. Given an estimate of α in the more general model, these two special cases can be readily tested using standard procedures such as a likelihood ratio test or an asymptotic t test. Of course, more general forms of homoskedasticity could be hypothesized. However, the form given here is parsimonious in terms of additional parameters to be estimated, maintains homogeneity of degree 0 in prices and expenditure, and contains the two special cases of interest described above.

4.2. Linear Expenditure System

The LES demand equations in share form (after deleting the observation subscript) are

$$w_i = \frac{p_i b_i}{\mu} + a_i\left(1 - \frac{\sum p_k b_k}{\mu}\right). \tag{34}$$

For this model allowing either the a's or the b's to be stochastic leads to computationally tractable disturbance terms in (34).

Consider first replacing the a_i in (34) by $a_i + \varepsilon_i$ where, as above, the ε_i's are normally distributed random variables with mean 0. The resulting disturbance in (34) defined as v_i, is

$$v_i = \varepsilon_i\left(1 - \frac{\sum p_k b_k}{\mu}\right), \qquad i = 1, \ldots, n. \tag{35}$$

The v_i's have mean 0 and covariance matrix Ω given by

$$\Omega_{ij} \equiv E(v_i v_j) = E(\varepsilon_i \varepsilon_j)\left(1 - \frac{\sum p_k b_k}{\mu}\right)^2. \tag{36}$$

If we assume that $E(\varepsilon_i \varepsilon_j) = \sigma_{ij}$ is a constant, then this is simply the standard model with heteroskedasticity. Such a model can easily be estimated, but if we wish to impose the condition that the share disturbances be homoskedastic with respect to expenditure, then we could assume that

$$E(\varepsilon_i \varepsilon_j) = \left[\frac{\mu}{\mu - \sum b_k p_k}\right]^2 \sigma_{ij}, \tag{37}$$

in which case $E(v_i v_j) = \sigma_{ij}$, which is the same as the standard stochastic specification.

If we replace the b_i's in (34) by $b_i + \varepsilon_i$, then the disturbance term in (34) becomes

$$v_i = \frac{p_i}{\mu}\varepsilon_i - a_i \sum \frac{p_k \varepsilon_k}{\mu}, \qquad i = 1, \ldots, n. \tag{38}$$

This differs from the corresponding expression for the QES, given by (28), only by the factor δ; hence, $\sum v_k = 0$ automatically. For reasons analogous to those given above for the QES, we assume that the u's satisfy

$$\text{(39)} \qquad E(\varepsilon_i \varepsilon_j) = \frac{\mu^2}{p_i p_j} \sigma_{ij}.$$

Because the b's enter the LES in essentially the same way that the c's enter the QES, this specification implies the same Ω matrix for the v's that appears in (31). After adjusting for heteroskedasticity and ensuring that the covariance matrix of share disturbances is homogeneous of degree 0 in expenditure and prices, the covariance matrix of disturbances is the same for the QES with random c's as it is for the LES with random b's.

Denoting this covariance matrix at observation t as Ω_t, the logarithm of the likelihood function for a random sample of T observations (aside from an additive constant) is then given by

$$\text{(40)} \qquad L(\beta, \Omega) = -\frac{1}{2} \sum_{t=1}^{T} \log|\Omega_t| - \frac{1}{2} \sum_{t=1}^{T} v_t' \Omega_t^{-1} v_t,$$

where $v_t' = (v_{1t}, \ldots, v_{n-1,t})$ is a $1 \times (n-1)$ vector of disturbances at time t, obtained from either the LES or QES share equations. Because Ω_t depends on t this likelihood cannot be concentrated with respect to the elements of Ω and hence must be maximized with respect to the σ's together with the demand system parameters. In this model, as in the standard one, the estimates are independent of which equation is dropped in the estimation. For a proof see Pollak and Wales [1969].

4.3. Translog Models

The basic translog model in share form (after deleting the observation subscript) is

$$\text{(41)} \qquad w_i = \frac{\alpha_i + \sum_j \beta_{ij} \log(p_j/\mu)}{\sum \alpha_k + \sum \sum_j \beta_{kj} \log(p_j/\mu)}, \qquad \begin{array}{l} \beta_{ij} = \beta_{ji} \quad \text{for all } i, j \\ \sum \alpha_k = 1. \end{array}$$

Suppose that we replace the α_i in (17) by $\alpha_i + \varepsilon_i$ where, as before, the ε_i are normal random disturbances; then both the numerator and denominator of (41) will be normally distributed. The distribution of a ratio of normally distributed variables was first investigated by Fieller [1932]. More recently Yatchew [1986] has extended Fieller's results to multivariate distributions involving ratios of normal variables. The density function for a set of ratios of normal variables that are not independent, with a common denominator, is in general very complex (Yatchew [1986]). Indeed, the density function has a closed-form representation only when there are an even number of ratios in the set (which, in the context of singular demand systems, implies an odd number of goods), or when there is only one ratio, in which case the density reduces to that derived by Fieller. Yatchew [1985] has estimated a Fieller distribution in the labor supply context, but as far as we know the more general case with more than one independent good has not been estimated, and is beyond the scope of this book.

A tractable special case of this model emerges when, instead of assuming that the ε's are independent, we assume that they satisfy $\sum \varepsilon_k = 0$ for each observation. In this case the ε's do not appear in the denominator and we have

$$E(v_i v_j) = \frac{\sigma_{ij}}{D^2}, \tag{42}$$

where D is the denominator in (41). Once again we have the standard model, but with a simple form of heteroskedasticity. To impose the condition that the share disturbances are homoskedastic we assume

$$E(\varepsilon_i \varepsilon_j) = D^2 \sigma_{ij}, \tag{43}$$

in which case $E(v_i v_j) = \sigma_{ij}$. Thus in this special case the random coefficient model reduces to the standard stochastic specification.

The empirical results of Section 5 of Chapter 6 are based on maximizing likelihood functions of the form given by (40) for the LES and QES demand systems.

6

Household Budget Data

In this chapter we report demand system estimates based on household budget data[1]. In Section 1 we demonstrate that interesting complete demand systems can be estimated from a small number of budget studies despite the limited price variability represented in such data. To illustrate we use household data for two years to estimate the LES and the QES. In Section 2 using a larger sample we concentrate on functional form specification and demographic influences on consumption patterns. We compare the LES with the quadratic expenditure system (QES), and the basic translog (BTL) with the generalized translog (GTL). We incorporate the age composition and the number of children using linear demographic translating and linear demographic scaling. In Section 3 we investigate three additional methods of incorporating demographic characteristics; because of the complexity of these procedures we estimate a simpler demand system—the generalized CES—rather than the QES or translog forms. In Section 4 we explore two dynamic QES specifications and introduce a stochastic structure that allows disturbances in each time period to be correlated across households. Due to the complexity of the estimation procedure we analyze separately households with different numbers of children. In Section 5 we explore random coefficient models for the LES and QES.

1. ESTIMATION FROM TWO BUDGET STUDIES

Both the LES and QES can be estimated using budget studies for two years. For any demand system, household budget data for a single period identify the income–consumption curve corresponding to the period's prices and, hence, the marginal budget shares at every level of expenditure. We consider first the LES, in which the demand equations in share form are given by

(1) $p_i x_i = p_i b_i + a_i(\mu - \sum p_k b_k), \qquad a_i > 0, \qquad (x_i - b_i) > 0, \qquad \sum a_k = 1.$

[1] This chapter is drawn in part from Pollak and Wales [1978, 1980, 1981] and Darrough, Pollak, and Wales [1983].

In the LES, the marginal budget shares are independent of prices and the level of expenditure and are equal to the a's. Thus, household budget data for a single period identify the a's. If one of the b's is known a priori, then budget data for a single period are enough to identify also the (n-1) remaining b's.[2] Even if none of the b's is known a priori, budget data for two periods identify all of the parameters of the LES: data from each period identify the corresponding income–consumption curve, and the intersection of the two linear income–consumption curves uniquely determines the point $(b_1, \ldots, b_n)$.

For the Σ-QES the demand equations in expenditure form are given by

$$(2) \quad p_i x_i = p_i b_i + a_i(\mu - \sum p_k b_k) + (p_i c_i - a_i \sum p_k c_k) \prod p_k^{-2a_k} (\mu - \sum p_k b_k)^2, \qquad \sum a_k = 1.$$

We now prefer this form to the λ-QES, which we estimated in Pollak and Wales [1978]. Budget data for two periods identify all of the parameters of the QES. The following heuristic argument, although not formally decisive, indicates why.[3] Data from each period identify the income–consumption curve corresponding to that period's prices, and since all income–consumption curves radiate upward from the point $(b_1, \ldots, b_n)$, the intersection of two income–consumption curves determines the point $(b_1, \ldots, b_n)$. Estimates of the b's enable us to calculate supernumerary expenditure $(\mu - \sum p_k b_k)$ for each household in each period. The a's are the coefficients of supernumerary expenditure, while the c's can be disentangled from the coefficients of $(\mu - \sum p_k b_k)^2$.

In contrast to the LES and QES, which can be estimated from two cross sections, in the translog demand system the number of budget studies needed to estimate all of the parameters varies with the number of goods (n).[4] In particular, it can be shown that M cross sections provide enough information to permit estimation of a translog system with at most $n + 1 + M(n - 1)$ parameters, while the nonhomothetic basic translog system contains $(n^2 + 3n - 2)/2$ independent parameters. Hence, two cross sections are sufficient to identify the translog with two or three goods, but not the translog with four or more goods; three cross sections identify it with four goods, but not with five or more.

---

[2] Howard Howe [1975] has shown that the identification of the parameters of Lluch's [1973] extended LES can be interpreted in precisely these terms: Lluch's specification implicitly assumes that the b associated with saving is 0.

[3] More formally, suppose that we rewrite (2) as $p_i x_i = \theta_{1i} + \theta_{2i}\mu + \theta_{3i}\mu^2$ and estimate the θ's for n − 1 goods in both periods. Then it is easy to show that after substituting the θ_{3i} values into the θ_{2i} values we can obtain estimates of the a's and of $\sum p_k b_k$ for both periods. Using these we can then obtain estimates of the individual b's from the θ_{1i} values, and estimates of the c's from the θ_{3i} values.

[4] See Lau, Lin, and Yotopoulos [1978] for estimates of a translog for Taiwan based on household budget data.

The consumption data used in this chapter are obtained from the *Family Expenditure Survey* series (U.K. Department of Employment and Productivity), an annual publication that reports expenditure patterns of households in the United Kingdom, cross-classified by income and family size.[5] In this section we use data for the two years 1966 and 1972; to simplify computations, we analyze only three consumption categories, "food," "clothing," and "miscellaneous."[6] For 1966 we have mean expenditure on each consumption category by families in four income classes and three family size classes: "one child," "two children," and "three or more children." For 1972, we have six income classes for families with one child and two children, five income classes for families with three children, and three income classes for families with "more than three children." These 32 cells form our basic data and we treat them as if they represent the consumption patterns of households rather than cell means.[7] Retail price indexes corresponding to these categories are taken from the *Annual Abstract of Statistics*, 1974.

We obtain a stochastic form for the LES and QES by adding a disturbance term to the share form of each demand equation. We use the share forms because they are likely to involve less heteroskedasticity than the expenditure forms. We demote the 3×1 vector of disturbances corresponding to the ith cell by $u_i = (u_{i1}, u_{i2}, u_{i3})$ and assume that $E(u_i) = 0$, that $E(u_i u_i') = \Omega$ for all i, and that the u_i's are independently normally distributed. This is the standard model discussed in Section 1 of Chapter 5; hence, the concentrated log likelihood is given by Eq. (3) in Chapter 5.

Estimates of the LES and the QES obtained in this way satisfy the regularity conditions required by economic theory. The Slutsky matrix implied by our parameter estimates is negative semidefinite at each of the

[5] A detailed discussion of the data used in this chapter appears in Appendix A.

[6] Since the purpose of this section is to illustrate demand estimation with as few as two cross sections, it is of little importance which years are selected provided they are far enough apart to provide some price variability. We choose 1966 and 1972 because we estimate systems that include the number of children as a determinant of consumption behavior and the period 1966–1972 is the longest period over which "children" were defined consistently. In the surveys before 1966 the number of persons rather than the number of children was reported. After 1972 persons were classified as children if they were 18 or under, while in earlier years they were so classified if they were 16 or under.

[7] Since our consumption data are cell means rather than observations on individual families, an aggregation problem arises. With the LES, it causes no difficulty because the mean consumption pattern of a group is the consumption pattern corresponding to the group's mean expenditure. But with the QES, the mean consumption pattern of a group depends on the variance as well as the mean of the group's expenditure. Unfortunately, we do not know these variances, so, faute de mieux, we have treated the reported cell means as if they were observations on individual families. Aggregation of demand systems quadratic in the independent variable is discussed by Klein [1962, pp. 25–26] and Diewert [1974b, pp. 129–130].

Table 1 Marginal Budget Shares and Own-Price Elasticities: LES and QES

Expenditure (S. per week)	Consumption category: Food	Clothing	Miscellaneous
	Marginal budget shares: QES		
200	.48	.25	.27
300	.39	.21	.40
400	.31	.17	.52
	Own-price elasticities: QES		
200	−.56	−.53	−.41
300	−.76	−1.16	−.90
400	−1.01	−1.80	−1.25
	Own-price elasticities: LES		
200	−.62	−1.06	−1.35
300	−.71	−1.04	−1.17
400	−.76	−1.02	−1.11

Notes:
1. Mean expenditure in 1970 is 315 S. per week.
2. For the LES the marginal budget shares are independent of expenditure and are .35, .20, and .45.

32 price expenditure situations corresponding to our data.[8] Further, the QES represents a significant improvement over the LES according to the usual likelihood ratio test; the 5% critical value is 7.8, while the value of $-2\log\lambda$ (where λ is the ratio of the likelihoods associated with the LES and the QES) is 10.4.

Perhaps the most interesting comparison to be made between the LES and the QES involves the behavior of the marginal budget shares. For the LES they are the estimated a_i parameters, while for the QES they depend on all of the estimated parameters. In Table 1 we present the marginal budget shares corresponding to 1970 prices for three expenditure levels.[9] For the QES the shares vary considerably with expenditure levels: at the 200 shilling expenditure level 48% of an additional shilling goes to food and at the 400 shilling level 31% of an additional shilling goes to food. For the LES the share is 35% at all expenditure levels.

[8]Our approach to the evaluation of regularity conditions differs from that adopted by Jorgenson and Lau [1979] in their discussion of the translog. They consider only whether regularity conditions are satisfied at a single point, the point of approximation of the translog, whereas we have checked them at every sample point. On the other hand, Jorgenson and Lau test the significance of the ability of their estimated translog to satisfy the regularity conditions, whereas we do not.

[9]This choice of year and expenditure levels makes possible comparison of these results with those of succeeding sections in this chapter.

Corresponding patterns for clothing and miscellaneous can be read from the table The own-price elasticities are also presented in Table 1 and range from $-.41$ to -1.80 for the QES, and from $-.62$ to -1.35 for the LES. In general the QES elasticities vary more with expenditure levels than do the LES elasticities, and the two sets of estimates suggest a different pattern of price responsiveness for the clothing category.

In summary we have estimated two complete demand systems—the LES and the QES—using U.K. household budget data for two periods. The empirical results are generally good in that the regularity conditions are satisfied at all sample points and the price and income responses appear to be reasonable, a priori.

The QES is well-suited to the analysis of data sets with severely limited price variation. With household budget data, each study identifies an income–consumption curve, and, although none of the underlying parameters of the QES can be identified from a single income–consumption curve, all of the parameters can be identified from two such curves.

2. FUNCTIONAL FORM COMPARISONS

We now use a time series of five budget studies to analyze the effects of demographic variables on expenditure patterns and to study the suitability of several different functional forms. We use the period 1968–1972 because this is the longest period over which the demographic variables are defined consistently.

We use the same consumption categories as in Section 1: food, clothing, and miscellaneous. The Family Expenditure Survey cross-classifies households by income and number of children, and for each cell it reports mean expenditure on the three consumption categories and the mean number of children under 2, the mean number under 5, and the mean number under 16. These 81 cells are our basic data and we treat them as if they represent the consumption patterns of households rather than cell means.[10]

The demand systems that we estimate are the basic translog (BTL), the generalized translog (GTL), and the quadratic expenditure system (QES).[11] In share form the demand equations for the BTL are

$$(3) \quad w_i = \frac{\alpha_i + \sum_j \beta_{ij} \log(p_j/\mu)}{\sum \alpha_k + \sum_j \sum \beta_{kj} \log(p_j/\mu)}, \qquad \beta_{ij} = \beta_{ji}, \qquad \sum \alpha_k + \sum \sum \beta_{kj} = 1,$$

[10] We describe the composition of these cells in Appendix A.

[11] We also estimated the LES and linear translog (LTL). However, likelihood ratio tests indicate that the QES is superior to the LES at the 1% level and that the GTL is superior to both the LES and the LTL at the 1% level.

where the α's and the β's are parameters to be estimated.[12] The GTL demand equations, in share form, are given by

$$(4)\qquad w_i = \frac{b_i p_i}{\mu} + [1 - (\sum p_k b_k)/\mu] \frac{\alpha_i + \sum_j \beta_{ij} \log[p_j/(\mu - \sum p_k b_k)]}{\sum \alpha_k + \sum \sum_j \beta_{kj} \log[p_j/(\mu - \sum p_k b_k)]},$$

where the α's, β's and b's are parameters to be estimated.

The Σ-QES demand equations in share form are given by

(5)

$$w_i = \frac{p_i b_i}{\mu} + \alpha_i\left(1 - \frac{\sum p_k b_k}{\mu}\right) + \left(\frac{p_i c_i}{\mu} - \alpha_i \sum \frac{p_k c_k}{\mu}\right) \prod \left(\frac{p_k}{\mu}\right)^{-2\alpha_k} \left(1 - \frac{\sum p_k b_k}{\mu}\right)^2,$$

$$\sum \alpha_k = 1.$$

We consider both linear demographic translating and linear demographic scaling as alternative procedures for introducing demographics into these demand systems. In the QES and GTL, introducing demographic variables by linear translating is equivalent to specifying that the translation parameters in these systems (the b's) are linear functions of the demographic variables.[13] The BTL contains no translation parameters; rather than apply demographic translating to a demand system without such constant terms, we modify the BTL by introducing them. Since this yields the GTL, our discussion of demographic translating involves only two distinct demand systems rather than three. Introducing demographic variables by linear scaling requires that each p_i on the right-hand side of the share equations be replaced by $p_i m_i$, where m_i is a linear function of the demographic variables.[14] For each cell, our data report the average number of children per household in each of three age categories: η_2, children under 2 years of age; η_3, children under 5 but at least 2 years of age; and η_4, children 5 and over but under 17. Hence for scaling we have $m_i = 1 + \bar{\varepsilon}_{2i}\eta_2 + \bar{\varepsilon}_{3i}\eta_3 + \bar{\varepsilon}_{4i}\eta_4$ and for translating we have $b_i = b_i^* + \bar{\delta}_{2i}\eta_2 + \bar{\delta}_{3i}\eta_3 + \bar{\delta}_{4i}\eta_4$. The average number of children in the

[12] This normalization on the β's and α's is the one used in Pollak and Wales [1980], on which this material is based, and differs from the one used in Chapter 2 ($\sum \alpha_k = 1$). The results are, of course, invariant to the choice of normalization.

[13] Some demand systems may be undefined or imply negative consumption of some goods for certain values of the b's. Fortunately, these problems did not arise in our sample.

[14] For the demand functions to make sense, all of the m's must be positive. Although our linear specification does not constrain the estimates of the m's to be positive, the m's implied by our estimates were positive at all sample points. However, our sample points are group averages, and the m's implied by our estimates are not positive for all possible family compositions. For example, our estimates of the QES imply a negative m for miscellaneous for a family with one child under two and another between two and five. If any of the m's are negative at sample points, then we must either reestimate the system after imposing nonnegativity constraints or use a different functional form for the m's; for example, the form $m_i = \eta^{\varepsilon_i}$ is guaranteed to yield positive values.

household, η_1, does not appear explicitly in this specification because it would be redundant. Since $\eta_1 = \eta_2 + \eta_3 + \eta_4$, we have three rather than four independent demographic variables, and a complete analysis can be based on any three of them. It is convenient to reformulate our specification of demographic effects to include η_1 instead of η_4. To do this, we replace η_4 by $\eta_1 - \eta_2 - \eta_3$ and substitute into the expressions for m_i and b_i. For m_i this yields $m_i = 1 + \varepsilon_{1i}\eta_1 + \varepsilon_{2i}\eta_2 + \varepsilon_{3i}\eta_3, i = 1, 2, 3$, where $\varepsilon_{1i} = \bar{\varepsilon}_{4i}$, $\varepsilon_{2i} = \bar{\varepsilon}_{2i} - \bar{\varepsilon}_{4i}$, and $\varepsilon_{3i} = \bar{\varepsilon}_{3i} - \bar{\varepsilon}_{4i}$. For b_i this yields $b_i = b_i^* + \delta_{1i}\eta_1 + \delta_{2i}\eta_2 + \delta_{3i}\eta_3$ where $\delta_{1i} = \bar{\delta}_{4i}, \delta_{2i} = \bar{\delta}_{2i} - \bar{\delta}_{4i}$, and $\delta_{3i} = \bar{\delta}_{3i} - \bar{\delta}_{4i}$.

For what parameter values are consumption patterns independent of household size or composition? In the scaling specification, if the ε's are all equal to zero, then demographic variables do not affect consumption patterns; if all the ε_2's and ε_3's are zero, while the ε_1's are not zero, then consumption patterns depend on household size, but not on household composition. In the translating specification, if all the δ's are zero, then household consumption patterns do not depend on the household's demographic characteristics. If all the δ_2's and δ_3's are zero, while the δ_1's are not zero, then consumption patterns depend on household size but not on household composition.

The number of parameters to be estimated depends on both the underlying functional form of the demand system and on the demographic specification. Take the case of three goods and three independent demographic characteristics. Scaling introduces nine additional parameters to be estimated regardless of the functional form of the demand system. Translating also introduces nine additional parameters to be estimated in the case of the QES and the GTL, demand systems that already include translation parameters. On the other hand in the BTL, a demand system that does not include translation parameters, demographic translating introduces twelve additional parameters.

As in Section 1 we obtain a stochastic specification for these demand systems by adding a disturbance term to the share form of each equation. We denote the 3×1 vector of disturbances corresponding to the ith cell by $u_i = (u_{i1}, u_{i2}, u_{i3})'$ and assume that $E(u_i) = 0$, that $E(u_i u_i') = \Omega$ for all i, and that the u_i are independently normally distributed. Since the dependent variables and the nonstochastic terms in the equations are shares and sum to 1 for each cell, the covariance matrix is singular. Hence, maximum liklihood estimates of the system can be obtained by minimizing the determinant of the sample error covariance matrix with respect to the parameters after dropping any equation.

In Table 2 we present the log likelihood values for the 13 models that we have estimated. We use these results as a basis for discussing three questions: first, do the demographic variables matter? In particular, are consumption patterns significantly affected by the number of children in the household, and, if so, do their ages have a significant additional effect? Second, which of the two methods of incorporating demographic

Table 2 Log Likelihood Values and Regularity Conditions

Model	Method of including variables	Number of parameters	Log likelihood	Regularity conditions
1. QES–no demographic variables	—	8	610.25	69
2. QES–number of children	T	11	696.35	71
3. QES–number of children	S	11	695.36	81
4. QES–number and age of children	T	17	705.69	76
5. QES–number and age of children	S	17	719.93	81
6. BTL–no demographic variables	—	8	597.86	81
7. BTL–number of children	S	11	685.05	81
8. BTL–number and age of children	S	17	706.79	80
9. GTL–no demographic variables	—	11	610.51	81
10. GTL–number of children	T	14	701.19	81
11. GTL–number of children	S	14	703.67	81
12. GTL–number and age of children	T	20	717.85	68
13. GTL–number and age of children	S	20	722.95	81

Notes:
1. T and S indicate whether demographic variables were incorporated through translating or scaling.
2. A common additive constant is omitted from each of the reported log likelihood values.
3. The regularity conditions column displays the number of sample points at which the estimated demand system is well-behaved. The total number of sample points is 81.

variables—translating or scaling—is more consistent with our data? Finally, which of the three demand systems we have estimated—the QES, the BTL, and the GTL—is most consistent with our data?

We consider first the question of whether the number of children and their age distribution affects the consumption patterns of the households in our sample. For all three demand systems, the addition of the number of children to the basic model is significant at the 1% level using both scaling and translating. Similarly the subsequent inclusion of the age distribution variables is significant at the 1% level in all models. Thus it appears that both the number and age of children have a significant effect on consumption patterns, even at the level of our three broad consumption categories, for all three functional forms and for both demographic specifications.

We next compare translating and scaling as alternative methods of incorporating demographic variables into demand systems. From Table 2 we see that scaling results in a higher likelihood value than translating in three of the four comparisons. However, since scaling and translating are not nested specifications, we cannot test the significance of the differences in the likelihood values.

Finally, we compare functional forms. We compare the BTL and the QES functional forms in terms of the values of the likelihood functions reported in Table 2. Since the QES and the BTL (but not the GLT) have the same number of parameters for the case of three goods, a direct comparison of the likelihood functions is of interest; but once again, since the functional forms are not nested, we cannot test the significance of any differences. A comparison of rows 1 and 6, 3 and 7, and 5 and 8 reveals that the QES yields a higher likelihood value than the BTL for all three cases.

We compare the GTL and the BTL using the likelihood ratio test. We find the GTL is a significant improvement over the BTL at the 1% level in all three cases (rows 6 and 9, 7 and 11, and 8 and 13). On the basis of these tests, we have eliminated the BTL from further consideration and have focused on the QES and GTL results.

No meaningful comparison of the QES and the GTL is possible. Since these demand systems are not nested, a formal test cannot be made on the basis of the likelihood values. An informal comparison based on the values of the likelihood function is indecisive: the GTL yields higher likelihood values than the QES, but since it involves three more parameters than the QES, no inference can be drawn from higher likelihood values.

The next set of questions involves the magnitude of the effects of demographic variables on consumption patterns. We first consider the

Table 3 Marginal Budget Shares by Family Size and Expenditure

	Consumption category/number of children								
	Food			Clothing			Miscellaneous		
Expenditure (S. per week)	1	2	3	1	2	3	1	2	3
				QES–Translating					
200	.35	.38	.41	.29	.31	.34	.36	.31	.26
300	.28	.30	.33	.22	.25	.27	.50	.45	.40
400	.20	.23	.25	.16	.18	.20	.64	.59	.55
				QES–Scaling					
200	.37	.40	.41	.29	.30	.31	.34	.31	.28
300	.27	.31	.34	.23	.24	.26	.51	.45	.40
400	.16	.22	.27	.16	.19	.20	.67	.59	.53
				GTL–Translating					
200	.37	.41	.45	.28	.30	.32	.36	.30	.24
300	.27	.31	.34	.22	.24	.26	.50	.45	.39
400	.18	.21	.25	.18	.19	.21	.64	.59	.54
				GTL–Scaling					
200	.38	.43	.48	.28	.30	.31	.33	.27	.21
300	.26	.30	.35	.23	.24	.26	.51	.45	.40
400	.16	.20	.24	.19	.20	.21	.65	.60	.55

Notes:
1. These results correspond to rows 2, 3, 10, and 11 in Table 2.
2. Marginal budget shares are evaluated at 1970 prices. Mean expenditure in 1970 is 315 S. per week.

effect of household size and composition on the marginal budget shares, and then their effects on own- and cross-price elasticities.

In Table 3 we present estimated marginal budget shares corresponding to three levels of family size and three expenditure levels for the GTL and QES systems with scaling and translating; the marginal budget shares are evaluated at 1970 prices.[15] In all four systems, the marginal budget shares for food and clothing increase with family size at all expenditure levels, while that for miscellaneous falls. In addition, the marginal budget shares for food and clothing fall as expenditure rises, while that for miscellaneous rises. Perhaps the most interesting conclusion to be drawn from the table is that the marginal budget share estimates are very similar for both demand specifications and for both methods of incorporating family size.

In Table 4 we report the marginal budget shares corresponding to different family sizes and age distributions for the four models reported in Table 3; the marginal budget shares are all evaluated at 1970 prices, with expenditure of 300 S. per week. For each model, the table reports the marginal budget shares for nine demographic profiles. Specifically, for families with one, two, or three children, we define three age distributions, A_1, A_2, A_3, so that the average age of the children in the family rises in moving from A_1 to A_2 and from A_2 to A_3.[16] In three of the four models, as the average age of the children in the household rises, the marginal budget shares for food and clothing increase (or remain the same), while the marginal budget share of miscellaneous decreases. However, the magnitudes of the marginal budget shares vary with the form of the demand system and the method of incorporating the demographic characteristics. In contrast, when only family size was included, the marginal budget share estimates were not sensitive to the functional form of the demand system or to the method used to incorporate family size.

In Table 5 we report the estimated own-price elasticities corresponding to different family size and age distributions for the four models reported in Tables 3 and 4; the format follows Table 4. As with the marginal budget shares, the own-price elasticities depend on both the form of the demand system and the method of incorporating the demographic characteristics. Each model tells a somewhat different story about own-price elasticities,

[15] These comparisons are based on equations that include family size but not the age variables, and hence correspond to rows 2, 3, 10, and 11 in Table 2.

[16] The construction of the A's is as follows. For each family size (η_1 equal to 1, 2, or 3 children) we set the average number of children between the ages of 2 and 5 (η_3) at its sample mean. We then allow the average number of children under 2 (η_2) to take on three values—its sample mean, and approximately its smallest and largest sample values. When η_2 takes on its largest value we denote the age distribution as A_1; when η_2 takes on its mean value we denote the age distribution as A_2; and when η_2 takes on its smallest value we denote the age distribution as A_3. Since this is done for each family size (η_1 equal to 1, 2, or 3 children) the values of η_2 and η_3 corresponding to the A_1, A_2, and A_3 distributions differ for each family size. The range of sample values for the ages of children and their means for each family size are reported in Appendix A.

Table 4 Marginal Budget Shares by Family Size and Age Distribution

Number of children	Food A_1	Food A_2	Food A_3	Clothing A_1	Clothing A_2	Clothing A_3	Miscellaneous A_1	Miscellaneous A_2	Miscellaneous A_3
				QES–Translating					
1	.20	.23	.27	.19	.22	.25	.61	.54	.48
2	.24	.27	.29	.22	.25	.28	.54	.48	.42
3	.29	.32	.35	.28	.30	.33	.44	.38	.32
				QES–Scaling					
1	.15	.20	.25	.26	.25	.23	.58	.55	.52
2	.23	.26	.29	.27	.25	.23	.50	.49	.49
3	.29	.31	.31	.28	.26	.24	.43	.43	.45
				GTL–Translating					
1	.11	.20	.30	.18	.22	.27	.71	.58	.42
2	.16	.24	.34	.20	.24	.28	.64	.52	.38
3	.29	.39	.41	.26	.30	.31	.45	.31	.28
				GTL–Scaling					
1	.19	.21	.24	.23	.23	.23	.58	.55	.53
2	.18	.22	.26	.24	.25	.25	.58	.53	.50
3	.32	.35	.38	.27	.27	.27	.41	.38	.35

Note:
1. Marginal budget shares are evaluated at 1970 prices with expenditure of 300 S. per week. Marginal budget shares may not add to unity due to rounding. These results correspond to rows 4, 5, 12, and 13 in Table 2.

although one consistent pattern is the increase (in absolute value) in the price elasticity for food as age increases. This occurs for both models and both methods of incorporating the demographics. Also, in almost all cases the elasticities increase as the number of children increases for both models and for both methods.

Finally we consider the extent to which the 13 models correspond to well-behaved preferences. Table 2 displays the number of sample points at which the Slutsky matrix is negative semidefinite. In 8 of the models regularity conditions hold at all 81 sample points and in all other models they hold at over 80% of the data points. The translog models overall appear to be slightly superior to the QES on these grounds, although if just the models that include age and number of children are considered, the QES is marginally better. Returning to the comparison between scaling and translating, note that in all four cases regularity conditions are satisfied for at least as many sample points with scaling as with translating.

In summary we have studied two topics in empirical demand analysis: the specification of functional forms for the demand equations and two procedures for incorporating demographic variables. We found that both

Table 5 Own-Price Elasticities by Family Size and Age Distribution

Number of children	Food A_1	Food A_2	Food A_3	Clothing A_1	Clothing A_2	Clothing A_3	Miscellaneous A_1	Miscellaneous A_2	Miscellaneous A_3
				QES–Translating					
1	−.21	−.24	−.37	−.22	−.24	−.54	−.63	−.56	−.73
2	−.25	−.36	−.54	−.25	−.53	−1.04	−.56	−.74	−1.14
3	−.49	−.72	−1.02	−.95	−1.71	−2.80	−1.11	−1.77	−2.64
				QES–Scaling					
1	−.53	−.62	−.69	−1.34	−1.43	−1.54	−.81	−.84	−.88
2	−.65	−.71	−.78	−1.47	−1.58	−1.75	−.88	−.92	−.98
3	−.68	−.72	−.79	−1.55	−1.68	−1.94	−.92	−.98	−1.05
				GTL–Translating					
1	−.81	−1.10	−1.42	−1.07	−1.23	−1.42	−1.00	−1.59	−2.25
2	−.92	−1.17	−1.47	−1.17	−1.32	−1.52	−1.34	−1.94	−2.60
3	−1.24	−1.55	−2.11	−1.46	−1.70	−2.21	−2.48	−3.34	−5.41
				GTL–Scaling					
1	−.76	−.78	−.81	−1.31	−1.23	−1.16	−1.32	−1.34	−1.34
2	−.80	−.84	−.87	−1.36	−1.31	−1.25	−1.56	−1.56	−1.55
3	−.89	−.91	−.93	−1.36	−1.30	−1.25	−1.43	−1.38	−1.31

Consumption category/age distribution.

Note: See Table 4.

the number and age of children in a family have a significant effect on consumption patterns, regardless of whether the underlying demand system is assumed to be QES or translog, and regardless of whether the demographic effects are assumed to operate through translating or through scaling.

Although demographic translating and demographic scaling are not nested specifications, some comparison of these two procedures is possible. In three of the four comparisons, scaling resulted in a higher likelihood value; and in all four cases, scaling did at least as well as translating in terms of regularity conditions.

We found the GTL significantly superior to both the linear translog and the BTL. This suggests that the translog forms previously discussed in the literature are more restrictive than has generally been recognized, and that more flexible responses to changes in expenditure are required, at least for that analysis of household budget data.

The QES and the translog specifications can be compared using the values of the likelihood functions, provided the specifications have the same number of parameters. Using this likelihood function criterion, the QES was superior to the translog in all of the cases we can compare. The highest likelihood values were attained by the GTL, but since this demand system contains three more parameters than the QES, we were unable to compare these systems.

3. DEMOGRAPHIC SPECIFICATIONS

In this section we illustrate the use of the five general procedures discussed in Chapter 3 for incorporating demographic variables in demand analysis. Due to the complexity of some of these procedures we estimate the generalized CES demand system rather than the QES or BTL. To simplify the analysis further we consider just one demographic variable, the number of children in the household. We assume that this variable enters the functions specifying the five procedures in a linear manner. The sample consists of data for the period 1966 through 1972 and yields 108 cells, which, as before, we treat as consumption patterns of households rather than as cell means.

The generalized CES demand equations (in share form) are given by

$$(6) \qquad w_i = \frac{p_i b_i}{\mu} + \frac{\alpha_i^c p_i^{1-c}}{\sum \alpha_k^c p_k^{1-c}}[1 - (\sum p_k b_k)/\mu], \qquad \sum \alpha_k = 1,$$

where the α's, b's, and c are parameters to be estimated. The parameter c is the elasticity of substitution between "supernumerary quantities," $(x_i - b_i)$. We use the same three consumption categories as in Sections 1 and 2; with three goods the generalized CES contains six independent parameters.

As usual we obtain a stochastic specification for this demand system by adding a disturbance term to the share form of each demand equation. We denote the 3×1 vector of disturbances corresponding to the ith cell by $u_i = (u_{i1}, u_{i2}, u_{i3})$ and assume $E(u_i) = 0$, $E(u_i u_i') = \Omega$ for all i, and that the u_i are independently normally distributed. Once again this is our standard stochastic formulation with log likelihood given by (3) in Chapter 5.

Table 6 presents log likelihood values and some other statistics for seven specifications: the five procedures described in Chapter 3 for which demand equations are given in Appendix B; the "pooled" specification that combines data from different family types and estimates a single demand system, implicitly assuming that consumption patterns are independent of demographic variables; and the "unpooled" specification that estimates three separate demand systems, one for each family type, implicitly assuming that demographic variables affect all demand system parameters.[17]

From Table 6 it is clear that family size significantly affects consumption patterns regardless of the method used to incorporate it. Comparing the

[17] In the "unpooled specification" the family types are households with one child, two children, and three or more children. Since data are available on households with exactly three children in one year only (1972), these data are used together with those for households with four or more children in 1972 to form the family type consisting of three or more children.

Table 6 Log Likelihood Values

	Translating	Scaling	Gorman	Reverse Gorman	Modified Prais–Houthakker	Unpooled	Pooled
Log likelihood	861.6262	863.4707	864.4165	864.5088	864.5157	870.5274	772.212
Estimate of v	1	0	1.35	1.32	n/a	n/a	n/a
Estimate of c	2.57	2.31	2.66	2.63	2.58	1.66, 4.00, 2.09	2.88
Number of estimated parameters	9	9	10	10	9	18	6
Chi-square	17.80	14.11	12.22	12.04	12.02	n/a	196.63

Notes:
1. n/a indicates not applicable.
2. For translating and scaling the value of v is not estimated but is set to 1 and 0, respectively.
3. For the unpooled estimates, the values for c correspond to families with 1 child, 2 children, and 3 or more children.
4. The chi-square value is calculated as minus twice the difference between 870.5274 and the log likelihood value in the column.

pooled results with those obtained using each of the five procedures, we find a very substantial decrease in the value of the likelihood in each case. Further, comparing each of the five procedures with the unpooled specification, the decrease in the likelihood function is significant at the 5% level only for demographic translating.[18] Because these results confirm the validity of incorporating demographic variables into the generalized CES by means of our five procedures we do not report marginal budget shares or price elasticities for the pooled or the unpooled specifications.

All five procedures yield estimated parameters that correspond to well-behaved preferences in all or virtually all of the 108 price–expenditure–demographic situations represented in our data.[19] Further, all yield similar values of c, the elasticity of substitution between supernumerary quantities: estimates range from 2.31 to 2.66.[20, 21]

The modified Prais–Houthakker procedure makes a very strong showing against the other four procedures on the basis of the value of the likelihood function: the 9-parameter modified Prais–Houthakker procedure yields a value of the likelihood function that is greater than that of the other two 9-parameter procedures (translating and scaling), as well as that of the two 10-parameter procedures (Gorman and reverse Gorman).[22] No formal classical tests are possible, however, because none

---

[18] The decrease is not significant for demographic translating at the 1% level.

[19] For translating, scaling, and the Gorman procedures, the Slutsky matrix was negative semidefinite in all situations; for the reverse Gorman and modified Prais–Houthakker procedures, it was negative semidefinite in 107 of 108 situations.

[20] As Table 6 shows, the unpooled estimates exhibit a wider range.

[21] For all five procedures, the generalized CES is significantly superior to the familiar LES at the 5% level. The generalized CES reduces to the LES when $c = 1$, and we have tested whether our estimated value of c differs significantly from 1. With the LES, the modified Prais–Houthakker procedure yields a likelihood value of 862.5833, while the other four procedures imply identical estimating equations and yield a likelihood value of 859.6875. It is easy to verify that scaling and translating are equivalent for the LES. Their equivalence to the Gorman and reverse Gorman procedures is less obvious but follows from the observation that ν is not identified when $c = 1$. To see this, consider Eq. (B3) in Appendix B: when $c = 1$, the marginal budget shares are constant and the α's and ν enter only in terms of the form $[1 + (1 - \nu)\alpha_i\eta]b_i + \nu\alpha_i\eta$. Rewriting these as $b_i + \alpha_i[\nu + b_i[\nu + b_i(1 - \nu)]\eta$ and denoting $\alpha_i[\nu + b_i(1 - \nu)]$ by π_i, it is clear that estimates of b_i and π_i, but not ν, can be obtained: all values of ν yield the same likelihood values, since as ν changes, the α_i's can adjust to give the same estimates of π_i.

[22] We also estimated the special case of the modified Prais–Houthakker procedure in which the specific scales are the same for all goods, and the even more restrictive special case in which their common value is such that children receive the same weight as adults. Since our basic units of observation are households with two adults and at least one child, this implies a coefficient of 0.5, so that a household with four children (i.e., six persons) spending 400 S. per week purchases twice as much of every good as a household with one child (i.e., three persons) spending 200 S. per week. The likelihood values for these models are 860.5457 and 814.2324, respectively; at the 5% level, these restrictions are rejected against the unpooled specification.

of the other four procedures is nested in the modified Prais–Houthakker procedure, nor is it nested in any of them.[23]

Scaling compares favorably with the remaining three procedures. It has a higher likelihood value than translating, which is also a 9-parameter procedure. It is not rejected at the 5% level against the unpooled specification nor against either the Gorman or reverse Gorman procedures.[24]

The reverse Gorman procedure is superior to the Gorman on the basis of the values of their likelihood functions. Since both procedures involve ten parameters, this provides a plausible basis for comparison, although no formal classical tests are possible.[25]

Translating is the weakest of the five procedures. It has a lower likelihood value than the modified Prais–Houthakker procedure and scaling, which also have nine parameters, and it is rejected at the 5% (but not at the 1%) level against the Gorman and reverse Gorman procedures.[26]

Table 7 presents marginal budget shares and own-price elasticities implied by our estimates of each of the five procedures for households with one, two, and three children. Price elasticities and marginal budget shares are evaluated at 1970 prices and an expenditure level that is close to the median expenditure for households in 1970.

The marginal budget share estimates implied by the five procedures are strikingly similar. The four procedures other than translating permit marginal budget shares to vary with family size; all show that the marginal budget share for food increases slightly with family size, that for clothing is almost constant, while that for miscellaneous decreases slightly. In the generalized CES, translating necessarily yields marginal budget shares that are independent of the number of children. The marginal budget

[23] Although techniques have been proposed for testing and comparing nonnested hypotheses (e.g., Pesaran and Deaton [1978]; Davidson and MacKinnon [1981]; Pollak and Wales [1991]), their application is beyond the scope of this book.

[24] Muellbauer [1977] rejects scaling against the unpooled specification as a method for incorporating household size into demand analysis. In addition to using a different functional form, Muellbauer uses data for a somewhat different period (1968–1973), includes households with no children, and uses ten consumption categories, some of which include consumer durables.

[25] Both the Gorman and reverse Gorman procedures yielded estimates of v outside the range [0, 1]. This does not violate regularity conditions, nor does it imply predicted responses to changes in the demographic variables very different from those implied by the other procedures. The likelihood functions that correspond to the Gorman and reverse Gorman procedures each had two maxima, a local maximum corresponding to a value of v below 0, and a global maximum corresponding to a value greater than 1. The values corresponding to the global maxima are given in Table 6. The estimated value of v at the local maximum for the Gorman (reverse Gorman) procedure is $-.60$ ($-.13$) and the corresponding value of the likelihood function is 864.3657 (864.1946).

[26] In Section 2 we found translating to be generally inferior to scaling for the QES and GTL, both of which are nonlinear in expenditure. Thus, it seems unlikely that translating performs poorly in Table 6 merely because the generalized CES is linear in expenditure.

Table 7 Marginal Budget Shares and Own-Price Elasticities by Family Size: Alternative Procedures for Incorporating Demographic Effects

	Consumption category/number of children								
	Food			Clothing			Miscellaneous		
Procedure	1	2	3	1	2	3	1	2	3
				Marginal budget shares					
Translating	.29	.29	.29	.23	.23	.23	.48	.48	.48
Scaling	.27	.29	.30	.23	.23	.23	.50	.48	.47
Gorman	.27	.29	.31	.23	.22	.23	.50	.49	.47
Reverse Gorman	.27	.29	.30	.23	.22	.23	.50	.49	.47
Modified Prais–Houthakker	.27	.29	.31	.22	.22	.23	.51	.49	.46
				Own-price elasticities					
Translating	−.78	−.75	−.72	−1.43	−1.49	−1.56	−1.41	−1.51	−1.65
Scaling	−.81	−.76	−.71	−1.64	−1.55	−1.44	−1.57	−1.56	−1.54
Gorman	−.73	−.75	−.76	−1.43	−1.50	−1.58	−1.42	−1.52	−1.66
Reverse Gorman	−.73	−.74	−.76	−1.44	−1.50	−1.57	−1.42	−1.52	−1.65
Modified Prais–Houthakker	−.78	−.75	−.72	−1.44	−1.51	−1.46	−1.53	−1.53	−1.53

Note:
1. Price elasticities are evaluated at 1970 prices (all equal to 100) and an expenditure level of 300 S. per week. Marginal budget shares are also evaluated at 1970 prices, but are independent of expenditure.

share estimates for translating are essentially identical to those implied by the other four procedures for two-child households.

A comparison of estimated own-price elasticities also suggests that the differences among procedures are small, especially for households with two children. For one-child households, the largest difference among procedures occurs for clothing, but this difference is only .21; for three-child families, the largest difference, also for clothing, is .14. We have calculated both own- and cross-price elasticities corresponding to expenditure levels of 200, 300, and 400 S. per week at 1970 prices, although we report only own-price elasticities at 300 S. per week. Estimated cross-price elasticities are very similar for all five procedures and the size of these differences decreases with expenditure: the largest difference between comparable elasticities at 300 S. per week is only .15.

In summary, we have estimated and compared five procedures for incorporating demographic variables into complete demand systems. Our comparisons are based on a generalized CES demand system into which we incorporated a single demographic variable, the number of children in the household. We have also estimated the pooled specification in which demographic variables have no effect on consumption patterns, and the unpooled specification in which they affect all demand system parameters.

We rejected the pooled specification against each of the five procedures, indicating that the number of children does affect consumption patterns independently of the procedure used. Of the five procedures only demographic translating could be rejected against the unpooled specification, indicating that the other four procedures are reasonably consistent with the data. These four procedures imply similar responses to changes in prices, total expenditure, and the number of children. Although no formal ranking of these procedures is possible, statistical tests can be applied to pairs of procedures that are nested, and the likelihood value provides a plausible basis for comparing procedures with the same number of independent parameters. Using these criteria, translating made the weakest showing while the two Gorman procedures were dominated by scaling, and the modified Prais–Houthakker was best of all.

4. DYNAMIC AND STOCHASTIC STRUCTURE

In this section we augment the QES model by introducing dynamic and stochastic structure. Our dynamic structure allows for the possibility of taste change by assuming that some utility function parameters are time-dependent. Our stochastic structure allows disturbances across households to be correlated in a given year. Unlike the stochastic specification used in preceding sections, the error components structure we use here is consistent with the possibility that, for example, in a particularly cold year most households consume more fuel than usual. We consider first the dynamic and then the stochastic specification.

We estimate four dynamic specifications: (a) a linear time trend specification in which some demand system parameters (the b's) vary linearly with time; (b) a lagged consumption specification in which some demand system parameters (again the b's) vary linearly with a variable representing past consumption; (c) a constant tastes specification in which all demand system parameters remain fixed; and (d) a model in which both (a) and (b) hold. The time trend specification is one of exogenous taste change and can be written as

$$b_{it} = b_i^* + \beta_i t. \tag{7}$$

The lagged consumption specification is one of endogenous taste change and can be written as

$$b_{it} = b_i^* + \beta_i z_{it-1}, \tag{8}$$

where z_{it-1} is a variable representing past consumption. As discussed in Chapter 4 the interpretation of the lagged consumption specification depends on the variable chosen to represent past consumption. For example, if z_{it-1} represents the household's own consumption of the ith good in period $t-1$, then the implied model is one of habit formation; if z_{it-1} represents other households' consumption of the ith good in period $t-1$, then the specification is one of interdependent preferences. Incorporating both the time trend and lagged consumption into the model yields

$$b_{it} = b_i^* + \beta_i t + \gamma_i z_{it-1}. \tag{9}$$

All of these models contain the constant tastes specification as a special case in which the β's (and/or γ's) are set to 0.

The usual "independent" stochastic specification adds disturbance terms to the share equations and assumes these disturbances are independent across households and over time. In addition to this independent specification we consider an error components structure in which the disturbance term is the sum of two components, one independent across households and over time, and the other a "time-specific effect" (TSE), which is the same for all households in a particular year.[27] The independent specification is thus a special case of the error components structure in which the TSE is absent. The TSE allows a positive correlation between the disturbances of different households in a particular year.

To formalize the error components model, we let u_{rt} denote the $(n-1) \times 1$ vector of disturbances added to $n-1$ of the equations in the demand system (1), where r denotes the household and t the time period. These disturbances are the sum of two components

$$u_{rt} = e_t + \varepsilon_{rt}, \qquad \begin{array}{l} r = 1, \ldots, q_t \\ t = 1, \ldots, T, \end{array} \tag{10}$$

where q_t is the number of households in period t and T the number of time periods. We do not assume that household r in period t is the same as household r in period τ nor even that there are the same number of households in every period. We do assume that the $(n-1) \times 1$ disturbance vector e_t and ε_{rt} are independently normally distributed with zero means and covariance matrices given by Γ and Δ, respectively, where Γ is positive semidefinite and Δ positive definite. The covariance matrix for the u's is thus given by

$$E(u_{rt}u_{s\tau}) = \begin{cases} \Gamma + \Delta, & r = s, \quad t = \tau \\ \Gamma & r \neq s, \quad t = \tau \\ 0 & \qquad\quad\;\; t \neq \tau. \end{cases} \tag{11}$$

Under these assumptions the log of the likelihood is (aside from an additive constant)

$$L = -\frac{1}{2}\sum_{t=1}^{T}\left\{\log|W_t| + (q_t - 1)\log|\Delta| + \sum_{s=1}^{q_t} u'_{st}\Delta^{-1}u_{st} + \sum_{s=1}^{q_t}\sum_{r=1}^{q_t} u'_{st}\frac{(W_t^{-1} - \Delta^{-1})}{q_t}u_{rt}\right\}, \tag{12}$$

[27] Since we do not have data on the same households in successive years, we do not discuss three-component structures involving household-specific disturbance terms.

where $W_t = \Delta + q_t\Gamma$.[28] Since Γ and Δ are $n-1 \times n-1$ symmetric matrices this stochastic specification contains $n(n-1)$ independent parameters.[29]

We estimate the same QES form as in Sections 1 and 2, using data for 1967–1972.[30] As mentioned earlier we estimate separate models for families with one, two, and three or more children. To represent past consumption we use the average consumption of all households in the sample in the previous year. Because cell sizes vary, it is a weighted average of cell means

$$z_{it} = \frac{\sum_{r=1}^{q_t} \eta_{rt} x_{irt}}{\sum_{r=1}^{q_t} \eta_{rt}}, \tag{13}$$

where x_{irt} is consumption of the ith good by income group r in period t and η_{rt} denotes the number of households in income group r in period t.[31]

Table 8 contains log likelihood values (aside from an additive constant) for the various models. Perhaps the most interesting result is the statistical insignficance of the TSE in all cases. The TSE involves an additional three parameters, while the 5% critical value for the chi-square distribution with three degrees of freedom is 7.81. In no case does twice the difference in the log likelihoods exceed this value. Indeed in six of the comparisons the TSE results in no increase in the likelihood function; in these cases the constraint that Γ be positive definite results in all three of the TSE parameters being set to zero in the likelihood maximization. We consider next the dynamic aspects of the model.

The significance of the dynamic effects depends on the number of children in the household. For households with one child the static model

[28] The derivation of this likelihood function appears in Chapter 5, while a more extensive analysis of this type of error components model appears in Magnus [1982].

[29] We impose the restriction that Δ be positive definite by writing it as the product of a lower triangular matrix (ΔL) and its transpose, and estimate the elements of ΔL rather than Δ. The same procedure is followed for Γ, thus giving elements of ΓL.

[30] One year is dropped to allow for lagged consumption.

[31] We ignore aggregation problems that may call for higher moments of lagged consumption. In a habit formation specification, the appropriate variable is the average lagged consumption of the households in the cell; in an interdependent preferences specification, the appropriate variable is the average lagged consumption of households assumed to influence those in the cell. The only data we have are those reported in the previous year's sample, and because the samples are not panels, these are different households. Using average consumption in the previous year's sample is compatible with an interdependent preferences specification in which tastes depend on the average past consumption of all households with the same demographic profile, regardless of their income or expenditure. The connection with a habit formation specification is rather loose.

An alternative but substantially more complicated approach would relate the preferences of households in each percentile of the income or expenditure distribution to the consumption of those occupying that position in the previous year.

Table 8 Log Likelihood Values for Dynamic QES Specifications

Number of children	Time-specific effect	Static	Time trend	Lagged consumption	Lagged consumption and time trend
1	No	241.97	242.04	242.31	243.42
	Yes	241.97*	242.04*	242.31*	243.42*
2	No	253.41	257.04	254.48	260.82
	Yes	254.58	259.49	256.48	260.83
3 or more	No	230.50	237.76	236.65	241.91
	Yes	232.61	238.23	236.65*	241.91*

Notes:
1. The asterisk indicates cases in which including the TSE resulted in no increase in the likelihood value, as discussed in the text.

is acceptable, with neither the time trend nor lagged consumption having a significant impact on consumption patterns. For households with two children neither the time trend nor lagged consumption is significant individually, but they are marginally significant jointly. That is, starting with the full model we reject the hypothesis of constant tastes at the 5% level. Households with three or more children exhibit the most significant dynamic behavior, with both the time trend and lagged consumption individually and jointly highly significant. It is not clear to us why the dynamic element becomes increasingly significant with family size.

The habit formation or interdependent preferences interpretation of the lagged consumption model requires positive β's. For households with three or more children all three β's are negative in the full model (and when the time trend is omitted), while for two-children households only the β for food is negative. It is tempting to interpret negative β's as reflecting partial stock adjustment for consumer durables rather than habit formation (see Houthakker and Haldi [1960]), although the absence of interest rates suggests that the interpretation is rather loose. The consumer durable interpretation would be more convincing if the β''s for clothing were negative for all family sizes.

We present marginal budget shares and own-price elasticities in Table 9 for our preferred models. For one-child families we base our results on the static model, while for the other families we base them on the full model containing both time-trend and lagged-consumption terms. The results are presented for 1970 with expenditure levels given by the sample points for that year.

Marginal budget shares follow the same pattern as a function of expenditure for all family sizes. The marginal budget share for food falls sharply, and that for miscellaneous rises sharply, as expenditure rises. The share of clothing falls sharply for all families except those with two children,

Table 9 Marginal Budget Shares and Own-Price Elasticities: Selected QES Specifications

	Marginal budget shares/ consumption category			Own-price elasticities/ consumption category		
Expenditure (S. per week)	Food	Clothing	Miscel-laneous	Food	Clothing	Miscel-laneous
Families with one child; static model						
220.0	.35	.26	.39	−.59	−1.08	−.72
242.2	.33	.25	.42	−.61	−1.14	−.79
284.8	.28	.23	.49	−.65	−1.26	−.88
399.5	.14	.17	.69	−.77	−1.61	−1.03
Families with two children; time trend and lagged consumption						
230.7	.40	.23	.37	−.74	−1.28	−1.00
259.0	.37	.23	.40	−.82	−1.46	−1.13
280.2	.35	.23	.42	−.87	−1.59	−1.22
320.2	.31	.23	.46	−.99	−1.83	−1.37
359.3	.27	.22	.51	−1.11	−2.07	−1.50
476.2	.16	.21	.63	−1.55	−2.78	−1.79
Families with three or more children; time trend and lagged consumption						
240.1	.42	.33	.25	−1.05	−3.09	−2.72
291.2	.38	.29	.33	−.58	−.94	−1.02
331.9	.34	.26	.40	−.37	−.37	−.52
480.6	.19	.15	.66	−.48	−.79	−1.17

for which it falls only slightly as income rises. These results are consistent with those in earlier sections.

The pattern of own-price elasticities as a function of expenditure depends on family size. These elasticities increase (in absolute value) as income rises for families with one and two children, but decrease for families with three or more children. The pattern for one- and two-child families is consistent with the results reported in Section 1 based on estimates that did not take family size into account.

Our estimated utility functions satisfy the regularity conditions of economic theory in almost all cases. For our preferred models the Slutsky matrix is negative semidefinite at all sample points for families with one or two children (a total of 64), and at 29 of 32 sample points for families with three or more children.

In summary we have explored two extensions to the models discussed in previous sections. First, we have allowed for the possibility of taste change, and find this significant for families with two or more children. Second, we have allowed disturbances across households to be correlated in a given year; but in no cases are these error components effects significant.

5. RANDOM COEFFICIENTS MODELS

In this section we consider random coefficients models for some of our functional forms. In Chapter 5 we showed that for the QES and LES a random coefficients stochastic structure implies share equation disturbances with a multivariate normal distribution; for these functional forms, such a stochastic structure is computationally tractable. For the BTL and GTL, on the other hand, the share equation disturbances resulting from a random coefficients specification involve ratios of normal variables; estimation of such models is beyond the scope of this book. In this section we use the same data set as in Section 2 to estimate a random coefficients model for the QES, with age and number of children incorporated either through scaling or translating. We also present results for the LES, although the restrictions imposed by the LES are clearly rejected by the data.

The QES demand equations in share form are given by (5) above. For reasons discussed in Chapter 5 we assume that the c's are randomly distributed across the sample. Hence the c_i's are replaced in (5) by $c_i + \varepsilon_i$ to give

$$(14) \quad w_i = \frac{p_i b_i}{\mu} + a_i\left(1 - \sum \frac{p_k b_k}{\mu}\right) + \left\{\frac{p_i}{\mu}(c_i + \varepsilon_i) - a_i \sum \frac{p_k}{\mu}(c_k + \varepsilon_k)\right\} \prod \left(\frac{p_k}{\mu}\right)^{-2a_k} \left(1 - \sum \frac{p_k b_k}{\mu}\right)^2.$$

We assume that the covariance matrix for the ε's satisfies

$$(15) \quad E(\varepsilon_i \varepsilon_j) = \frac{\sigma_{ij}\mu^{2\alpha}}{p_i^\alpha p_j^\alpha \delta^2(P, \mu)}, \qquad i = j$$
$$= 0, \qquad i \neq j,$$

where $\delta(P, \mu) \equiv \prod \left(\frac{p_k}{\mu}\right)^{-2a_k} \left[1 - \sum \frac{p_k b_k}{\mu}\right]^2$ and α is a parameter to be estimated. Further the ε's are assumed to be normally distributed and uncorrelated across observations, and the σ_{ij} are constants. Denoting the disturbance for the ith share equation at observation t in (14) as v_{it} and the corresponding disturbance covariance matrix as Ω_t, the ijth element of Ω_t is given by

$$(16) \quad E(v_{it} v_{jt}) = \frac{p_i p_j \sigma_{ij} - p_i a_j \sum p_k \sigma_{ki} - a_i p_j \sum p_k \sigma_{kj} + a_i a_j \sum \sum p_k p_\ell \sigma_{k\ell}}{p_i p_j},$$

where for simplicity the t subscripts have been dropped from the p's.

The LES demand equations in share form are given by (14) above except that the last term on the right-hand side is set to zero. We assume that the b's are randomly distributed and replace the b_i by $b_i + \varepsilon_i$ to give

$$w_i = \frac{p_i}{\mu}(b_i + \varepsilon_i) + a_i\{1 - \sum \frac{p_k}{\mu}(b_k + \varepsilon_k)\}. \tag{17}$$

If we assume

$$E(\varepsilon_i \varepsilon_j) = \frac{\sigma_{ij}\mu^{2\alpha}}{p_i^\alpha p_j^\alpha}, \qquad i = j \tag{18}$$
$$= 0, \qquad i \neq j$$

and retain the other assumptions about the ε's that we made for the QES, then the disturbance covariance matrix for (17) will again be Ω_t as defined in (16). These covariance matrices are the same because, as we noted in Chapter 5, the b's enter the LES in essentially the same way that the c's enter the QES.

The log likelihood function for a sample of T observations (aside from an additive constant) is then given by (40) in Chapter 5 for both the LES and QES. For simplicity we assume that $\sigma_{ij} = 0$ for $i \neq j$. Since Ω depends on t, the likelihood function cannot be concentrated with respect to these σ_{ii} parameters. As discussed in Chapter 5 this introduces n covariance parameters (σ_{ii}) to be estimated; with $n = 3$ this is the same number of covariance parameters as appears in the standard additive model.[32]

We incorporate the age and number of children into the nonstochastic part of (14) through the use of demographic translating and demographic scaling as in Section 2.

In Table 10 we present the log likelihood values and number of sample points (out of 81) at which the estimated demand systems satisfy the regularity conditions. A number of interesting conclusions can be drawn. First, using standard likelihood ratio tests in all cases we find the value of α significantly different from zero, but not significantly different from one. This corresponds with our a priori beliefs, as discussed in Chapter 5, that the disturbances associated with the share equations will be homoskedastic as a function of income, whereas those associated with the expenditure equations will not be. In view of this finding we consider below only the random coefficients model based on $\alpha = 1$. A comparison of this model with the standard additive disturbance model reveals likelihood values that are virtually identical in all three cases, as indeed are the number of sample points satisfying regularity conditions. Further, although not reported here, the simultaneous inclusion of an additive disturbance and random coefficients disturbance never results in a significantly higher likelihood value than is obtained with the inclusion of

[32] The standard additive model includes $n(n-1)/2$ covariance parameters. We have also estimated the more general random coefficients models in which $\sigma_{ij} \neq 0$, $i \neq j$, but we do not report the results below since in no case are these additional parameters sgnificant.

Table 10 Log Likelihood Values and Regularity Conditions

	Log likelihood	Regularity conditions
	QES	
Demographic scaling		
Additive disturbance model	710.93	81
Random coefficients model		
$\alpha = 1$	710.25	81
$\alpha = 0$	707.20	81
α estimated (.64)	711.51	81
Demographic translating		
Additive disturbance model	705.69	76
Random coefficients model		
$\alpha = 1$	705.56	77
$\alpha = 0$	697.61	77
α estimated (.85)	705.81	78
	LES	
Additive disturbance model	671.66	81
Random coefficients model		
$\alpha = 1$	671.25	81
$\alpha = 0$	663.89	81
α estimated (1.003)	671.25	81

Notes:
1. The regularity conditions column records the number of sample points (out of 81) at which estimated demand systems satisfy the regularity conditions.
2. For the LES, translating and scaling are equivalent procedures.

just one type of disturbance. This suggests that there is little to choose between the two methods of incorporating the random component, and no need to use both.

In Table 11 we report marginal budget shares by family size and age composition for the additive disturbance and random coefficients models. The additive disturbance results for the QES are the same as those reported in Table 4 and are repeated here for ease of comparison with the random coefficients model. As is clear from the table, the two methods yield identical estimates in all cases (to two decimal places). Similar results hold for the own-price elasticities and are not presented here, since for the random coefficients model they are virtually identical to those given in Table 5 for the additive disturbance model.

In summary we have estimated random coefficients models for the LES and QES. These models are computationally tractable if the b's in the LES or the c's in the QES are assumed to be stochastic. Reasonable assumptions about the behavior of the share equation error variances as functions of income and prices yield the same disturbance covariance

Table 11 Marginal Budget Shares by Family Size and Age Distribution: Additive Disturbance and Random Coefficients Models

	Consumption category/age distribution								
	Food			Clothing			Miscellaneous		
Number of children	A_1	A_2	A_3	A_1	A_2	A_3	A_1	A_2	A_3
	QES-Scaling—Additive Disturbance Model								
1	.15	.20	.25	.26	.25	.23	.58	.55	.52
2	.23	.26	.29	.27	.25	.23	.50	.49	.49
3	.29	.31	.31	.28	.26	.24	.43	.43	.45
	QES-Scaling—Random Coefficients Model								
1	.15	.20	.25	.26	.25	.23	.58	.55	.52
2	.23	.26	.29	.27	.25	.23	.50	.49	.49
3	.29	.31	.31	.28	.26	.24	.43	.43	.45
	QES-Translating—Additive Disturbance Model								
1	.20	.23	.27	.19	.22	.25	.61	.54	.48
2	.24	.27	.29	.22	.25	.28	.54	.48	.42
3	.29	.32	.35	.28	.30	.33	.44	.38	.32
	QES-Translating—Random Coefficients Model								
1	.20	.23	.27	.19	.22	.25	.61	.54	.48
2	.24	.27	.29	.22	.25	.28	.54	.48	.42
3	.29	.32	.35	.28	.30	.33	.44	.38	.32

Notes:
1. See note to Table 4.
2. For the LES the marginal budget shares are constant and are .25, .24, and .51 for food, clothing, and miscellaneous for both the additive error model and the random coefficients model.

matrix in both of these cases. There are two major findings: first, the share but not the expenditure equation disturbances are homoskedastic with respect to income; and second, the random coefficients model yields virtually the same results as does the additive disturbance model.

APPENDIX A: DATA

As mentioned in the text, we analyze three broad consumption categories in this chapter: food, clothing, and miscellaneous. The categories are defined as follows. The food category does not include alcoholic drink. Clothing is officially "clothing and footwear." Miscellaneous is the sum of two categories from the survey, "other goods" and "services." The principal subcategories of other goods are "leather, travel and sports goods; jewelry; fancy goods, etc.," "books, magazines and periodicals," "toys and stationery goods, etc.," and "matches, soap, cleaning materials, etc." The principal subcategories of services are "radio and television, licenses and rentals," "educational and training expenses," and "subscriptions and donations;

hotel and holiday expenses; miscellaneous other services." The survey reports seven major expenditure categories that we have omitted entirely: "housing," "fuel, light, and power," "alcoholic drink,"tobacco," "durable household goods," "transport and vehicles," and "miscellaneous." Our three categories of food, clothing, and miscellaneous account for between 46 and 58% of total consumption expenditures.

Mean expenditures on the three consumption categories are available by income level and number of children (on a consistent basis) for the period 1966 through 1972. This yields 108 cells, which we used as basic data points in the equation estimations. In Section 1 only the two years 1966 and 1977 were used; in Sections 3 and 4, the seven years 1966–1972 were used. In Sections 2 and 5, the five years 1968–1972 were used since information on the age distribution of children was not available for 1966 and 1967. A tabulation of the number of cells by year, number of children, and number of income levels is given below.

Classification of Basic Data Cells

Year	Number of children in household	Number of income levels	Total number of cells
1966	1, 2, 3 or more	4	12
1967	1, 2, 3 or more	5	15
1968	1, 2, 3 or more	5	15
1969	1,2	5	10
	3 or more	4	4
1970	1,3 or more	4	8
	2	6	6
1971	1, 2, 3 or more	6	18
1972	1,2	6	12
	3	5	5
	4 or more	3	3
			108

The range and sample mean for the age variables by number of children in the household are as follows (for years 1968–1972):

Number of children in household	Age under 2	Age 2–4	Age 5–16
1	.10–.51 (.33)	.11–.25 (.20)	.31– .65 (.47)
2	.11–.55 (.30)	.32–.83 (.60)	.62–1.57 (1.10)
3	.10–.50 (.29)	.47–.81 (.63)	.69–2.41 (2.08)

For the results in Tables 4 and 5 we have set the "2–4" variable (η_3) at its mean of .20 and the "under 2" variable (η_2) at .16, .33, and .50, for a one-child household. For a two-child household the corresponding values are .60 for η_3 and .15, .30, and .45 for η_2; for a three-child household the values are .63 for η_3 and .15, .29, and .45 for η_2.

All households consist of one man, one woman, and at least one child.

The price indexes for food and clothing are taken directly from Table 4.9 of the 1974 *Annual Abstract of Statistics* (U.K. Central Statistical Office). The price index for the miscellaneous category is a weighted average of the "miscellaneous goods" and "services" indexes in the table, with the weights given by the table entries for 1966. Price indexes for all three categories are set to unity in 1970 and are given below.

Price Indexes

Year	Food	Clothing	Miscellaneous
1966	.825	.888	.787
1967	.846	.902	.811
1968	.879	.916	.866
1969	.935	.951	.926
1970	1.000	1.000	1.000
1971	1.111	1.068	1.109
1972	1.209	1.145	1.170

In Pollak and Wales [1978], we used price indexes obtained from Tables 29 and 30 of *National Income and Expenditure 1964–74* (U.K. Central Statistical Office, 1975). We were unaware of the existence of the retail price series at that time. For this reason and because we are now using the Σ-QES, the results reported in Section 1 differ from those reported in Pollak and Wales [1978].

APPENDIX B: GENERALIZED CES DEMAND SYSTEMS WITH DEMOGRAPHIC VARIABLES

Introduction of the number of children in the household (η) into the generalized CES demand system (6) using the five methods discussed in Section 3 gives the following equations:

Translating

$$\text{(B1)} \qquad w_i = \frac{p_i}{\mu}(b_i + \alpha_i \eta) + \frac{\alpha_i^c p_i^{1-c}}{\sum \alpha_k^c p_k^{1-c}}\left[1 - \frac{\sum p_k(b_k + \alpha_k \eta)}{\mu}\right].$$

Scaling

$$\text{(B2)}\quad w_i = \frac{p_i}{\mu}(1+\alpha_i\eta)b_i + \frac{\alpha_i^c[(1+\alpha_i\eta)p_i]^{1-c}}{\sum \alpha_k^c[(1+\alpha_k\eta)p_k]^{1-c}}\left[1 - \frac{\sum p_k(1+\alpha_k\eta)b_k}{\mu}\right].$$

Gorman

$$\text{(B3)}\quad w_i = \frac{p_i}{\mu}\{[1+(1-\nu)\alpha_i\eta]b_i + \nu\alpha_i\eta\} + \frac{\alpha_i^c\{[1+(1-\nu)\alpha_i\eta]p_i\}^{1-c}}{\sum \alpha_k^c\{[1+(1-\nu)\alpha_k\eta]p_k\}^{1-c}}$$
$$\times\left\{1 - \frac{\sum p_k\{[1+(1-\nu)\alpha_k\eta]b_k + \nu\alpha_k\eta]\}}{\mu}\right\}.$$

Reverse Gorman

$$\text{(B4)}\quad w_i = \frac{p_i}{\mu}[1+(1-\nu)\alpha_i\eta](b_i + \nu\alpha_i\eta) + \frac{\alpha_i^c\{[1+(1-\nu)\alpha_i\eta]p_i\}^{1-c}}{\sum \alpha_k^c\{[1+(1-\nu)\alpha_k\eta]p_k\}^{1-c}}$$
$$\times\left\{1 - \frac{\sum p_k[1+(1-\nu)\alpha_k\eta](b_k + \nu\alpha_k\eta)}{\mu}\right\}.$$

Modified Prais–Houthakker

$$\text{(B5)}\quad w_i = \frac{p_i}{\mu}b_i(1+\alpha_i\eta) + \frac{\alpha_i^c(1+\alpha_i\eta)p_i^{1-c}}{\sum \alpha_k^c(1+\alpha_k\eta)p_k^{1-c}}\left[1 - \frac{\sum p_k b_k(1+\alpha_k\eta)}{\mu}\right].$$

Note: All summations are over k from 1 to 3.

7

Per Capita Time-Series Data

In this chapter we use aggregate time-series data to study several issues in demand system estimation.[1] In Section 1 we investigate alternative demand system specifications and alternative estimation procedures. In Section 2 we examine the effects of pooling data from different countries for demand system estimation.

Section 1 begins by examining alternative specifications of functional form, of dynamic structure, and of stochastic structure. The functional forms we consider are those estimated using household budget data in Chapter 6, Section 2—the LES, QES, BTL, and GTL. The dynamic structures we consider are "dynamic translating" and "dynamic scaling." Both are general procedures in that they do not require the demand system to have a particular functional form, but can be used in conjunction with any complete demand system. The stochastic specifications we consider are the usual "independent" specification in which disturbances in different periods are unrelated, and two specifications permitting first order serial correlation: a "diagonal" specification allowing only one serial correlation parameter for the entire demand system, and a more general "free" specification involving additional serial correlation parameters.

Section 1 concludes with an investigation of estimation procedures that differ in their treatment of the first observation. We estimate each demand system in two ways: using a generalized first difference procedure that excludes the first observation, and using a maximum likelihood procedure that includes it. Our estimates in Section 1 are based on per capita time-series data for the U.S. for the period 1947 through 1983.

In Section 2 we explore the possibility of pooling data from different countries in the estimation of demand systems. Since pooling is most plausible for countries at similar stages of development, we use per capita time series data from three advanced industrial countries: Belgium, the U.K., and the U.S. A nonparametric revealed preference test indicates that the data from these three countries could not have arisen from maximizing a single static nonstochastic utility function. However, pooling may still be possible if there are some short-run or long-run differences in utility

[1] This chapter is drawn in part from Pollak and Wales [1987, 1992].

functions among countries. We estimate the QES under alternative specifications that permit such differences in an attempt to confirm the validity of pooling. We use purchasing power parities to transform the data before estimation.

1. ALTERNATIVE DEMAND SPECIFICATIONS AND ESTIMATION METHODS

The four demand systems we estimate are forms of the QES or the GTL. Since these were presented in Chapter 6, Section 2 we do not reproduce them here. The QES in share form is given by Eq. (5) in Chapter 6 and the GTL by Eq. (4). The LES and BTL are special cases of these.

1.1. Dynamic Structure

We estimate the usual static model, in which all demand system parameters remain fixed, and two dynamic specifications that permit some demand system parameters to vary with past consumption. As discussed in Chapter 4, the two dynamic specifications, dynamic translating and dynamic scaling, are general procedures for allowing systematic parameter variation in empirical demand analysis. That is, either specification can be used in conjunction with any original demand system, not just with a restricted class of functional forms. We describe both specifications as modifications of an original class of demand systems, $\{x_i = \bar{h}^i(P, \mu), i = 1, \ldots, n\}$. We assume these original demand systems are "theoretically plausible" (i.e., they can be derived from "well-behaved" preferences), and we denote the corresponding direct utility function by $\bar{U}(X)$ and the indirect utility function by $\bar{\Psi}(P, \mu)$.

Our dynamic translating procedure varies according to whether the original demand system contains constant terms; when it does, we assume that they depend on past consumption. In principle they may depend on any variables representing past consumption (e.g., a weighted average of all past consumption or the highest level of consumption attained in the past), but we assume that they depend only on the previous period's consumption. Denoting the constant terms by d's, we write $d_i = D^i(x_{it-1})$. Log linear dynamic translating, the specification we estimate, is given by

$$D^i(x_{it-1}) = d_i^* x_{it-1}^{\gamma_i} \tag{1}$$

and adds n parameters to the modified system.

When the original demand system does not contain constant terms, our dynamic translating procedure replaces the original system by

$$h^i(P, \mu) = d_i + \bar{h}^i(P, \mu - \sum p_k d_k). \tag{2}$$

This introduces n constant terms into the demand system, and we assume they depend on past consumption, as in the preceding case.[2] If the original demand system is theoretically plausible, then the modified system is also, at least for d's close to zero. The modified system satisfies the first order conditions corresponding to the indirect utility function $\Psi(P, \mu) = \bar{\psi}(P, \mu - \sum p_k d_k)$.

Dynamic scaling replaces the original demand system by

$$h^i(P, \mu) = m_i \bar{h}^i(p_1 m_1, \ldots, p_n m_n, \mu), \tag{3}$$

where the m's are scaling parameters that depend on the previous period's consumption, $m_i = M^i(x_{it-1})$.[3] If the original demand system is theoretically plausible, then the modified system is also, at least for m's close to one. The modified system satisfies the first order conditions corresponding to the indirect utility function $\Psi(P, \mu) = \bar{\Psi}(p_1 m_1, \ldots, p_n m_n, \mu)$ and the direct utility function $U(X) = \bar{U}(x_1/m_1, \ldots, x_n/m_n)$. Loosely speaking, we might interpret x_i/m_i as a measure of x_i in "efficiency units" rather than physical units.

Log linear dynamic scaling, the specification we estimate, is given by

$$M^i(x_{it-1}) = x_{it-1}^{\gamma_i}. \tag{4}$$

This specification guarantees that the implied value of m_i will be positive, as is required on theoretical grounds. It adds n parameters to the original demand system.

Specifications allowing demand system parameters to depend on past consumption are formally analogous to those allowing them to depend on demographic variables. Thus, dynamic translating is analogous to demographic translating and dynamic scaling to demographic scaling. Although we do not consider more general dynamic specifications here, all the demographic specifications discussed in Chapter 6, Section 3 have obvious dynamic counterparts.

1.2. Stochastic Structure

We estimate three stochastic specifications, one in which disturbances in different periods are independent and two others that allow first order serial correlation: the "diagonal" autocorrelation specification involves only one serial correlation parameter, while the "free" specification involves $(n-1)^2$ independent serial correlation parameters.

All of our stochastic specifications postulate an additive disturbance term on the share equations. We use the same notation as in Chapter 5

[2] A more general formulation of translating would replace (1) by $D^i(x_{it-1}) = c_i + d_i^* x_{it-1}^{\gamma_i}$. Given the computational difficulties in converging our simpler models, it is unlikely that this more general specification would converge.

[3] A more general formulation would allow m_i to depend on consumption in all previous periods.

and write

$$w_t = \omega(z_t, \beta) + \tilde{u}_t, \tag{5}$$

where $\tilde{u}_t = (u_{t1}, \ldots, u_{tn})'$, and $E(\tilde{u}_t) = 0$ for all t.

The free stochastic specification assumes that $\tilde{u}_t$ follows a general first order autoregressive process

$$\tilde{u}_t = \tilde{R}\tilde{u}_{t-1} + \tilde{e}_t, \tag{6}$$

where $\tilde{R}$ is an $n \times n$ matrix of autoregressive parameters and $\tilde{e}_t$ is an $n \times 1$ vector of disturbances with $E(\tilde{e}_t) = 0$ and $E(\tilde{e}_t\tilde{e}_t') = \Omega$ for all t.[4] The diagonal specification is a special case in which $\tilde{R}$ is diagonal; the independent specification corresponds to $\tilde{R} = 0$. Although the diagonal specification might appear to allow n independent serial correlation parameters, it allows only one (see Chapter 5). As usual we can estimate the demand systems after dropping one equation. With three goods, we can identify four independent transformations of the parameters of the $\tilde{R}$ matrix. We denote the 2×2 transformed $\tilde{R}$ matrix by R.[5]

1.3. Data and Estimation

Our estimates are based on annual U.S. per capita data for the years 1948–1983. We use three broad commodity groups constructed from the national product accounts: "food," "clothing," and a "miscellaneous" category that excludes shelter, consumer durables, and nondurables that seem closely related to them.[6] Excluding shelter and durables is justified if our three included commodity groups are separable from the excluded ones. The alternative of including the flow of services of shelter and consumer durables would require time series data representing these service flows and the corresponding implicit rents, and would require that consumption of these services be in equilibrium given their implicit rents.[7] Our three included commodity groups account for about half of "personal consumption expenditure." Precise definitions of our commodity groups are given in Appendix A.

---

[4]The only paper we know that estimates demand systems using a free R* matrix is Anderson and Blundell [1982]. They estimate a general dynamic short-run model, using a generalized first difference procedure.

[5]Recall from Chapter 5 that R is obtained from $\tilde{R}$ by first subtracting the last column of $\tilde{R}$ from each of the other columns of $\tilde{R}$, and then deleting the last row and column.

[6]The treatment of shelter is problematic. Although national product accounts purport to measure the flow of housing services, we have excluded shelter because we are skeptical about these data. Howe, Pollak, and Wales [1979] report estimates based on four commodity groups, including shelter.

[7]As an example of this approach see Diewert [1974a]. It is difficult to think of assumptions that justify ignoring durability and treating reported purchases of durables as current consumption flows.

There is evidence, especially strong in the single-equation context, that parameter estimates and hypothesis tests can be sensitive to omission of the first observation. The literature suggests that this is especially likely to be true in small samples when the independent variables exhibit time trends.[8,9] To investigate this phenomenon in the demand system context, we use two alternative estimation procedures discussed in Chapter 5: the "generalized first difference" procedure, which drops the first observation, and the "maximum likelihood" procedure, which does not. Because the generalized first difference procedure is asymptotically equivalent to maximum likelihood, however, the two procedures yield different results only in "small" samples.

1.4. Results

We report results for 30 different models: (i) three dynamic structures (static, dynamic translating, dynamic scaling), (ii) four functional forms (LES, QES, BTL, GTL), and (iii) three stochastic specifications (R = 0, R

Table 1 Number of Independent Parameters

	LES	QES	BTL	GTL
Static				
R = 0	5	8	8	11
R diagonal	6	9	9	12
R free	9	12	12	15
Dynamic translating				
R = 0	8	11	+	14
R diagonal	9	12	+	15
R free	12	15	+	18
Dynamic scaling				
R = 0	8	11	11	14
R diagonal	9	12	12	15
R free	12	15	15	18

Notes:
1. We do not include the disturbance covariance parameters in the count.
2. The + denotes the fact that translating can be applied only to systems containing translating parameters; when translating parameters are introduced into the BTL it becomes the GTL.

[8] In our sample the normalized prices for food and clothing exhibit a strong time trend, falling almost monotonically over the period. For the miscellaneous category, normalized prices rise initially until about the middle of the period and then fall.

[9] See, for example, Park and Mitchell [1980] andMaeshiro [1979] for single-equation estimation, and Beach and MacKinnon [1979] for systems estimation.

diagonal, R free).[10] Table 1 shows the number of independent parameters (other than disturbance covariance parameters) in each model. Table 2a presents log likelihood values (aside from an additive constant) for 30 models estimated with the generalized first difference procedure and shows the number of sample price–expenditure situations at which each estimated demand system corresponds to well-behaved preferences.[11] Table 2b presents the corresponding values for the maximum likelihood procedure. The likelihood values reported in Tables 2a and 2b are not comparable because the estimates are based on different procedures and different samples.[12]

The results of Tables 2a and 2b may be analyzed in a number of ways. We start with the most general models and then ask whether dynamic structure or stochastic specification restrictions corresponding to a simpler functional form can be imposed.[13] We begin with the LES–QES estimates; the most general models are the QES with R free, and either dynamic translating or scaling. We find that in none of the four such cases can the restrictions corresponding to the LES be imposed, and thus we conclude that the QES is a significant improvement over the LES. We find that in only one of the four cases—dynamic translating with the ML estimates—can the restrictions corresponding to a single serial correlation coefficient be imposed. This is also the only case in which the static model is accepted, and it is accepted only when R is free; when R is constrained to be diagonal, the static model is rejected. Thus, for three of our four most general models the restrictions corresponding to the LES, to a static model, or to a diagonal R are all rejected. In the fourth case—dynamic translating under ML—we reject the LES and either the static model or R diagonal, although not both.

With regard to dynamic specification, we find little to choose between translating and scaling. In our most general models translating yields a higher likelihood value than scaling with the first differencing estimates, but a lower value with the ML estimates. The first differencing estimates, however, do much better at satisfying regularity conditions than do the ML estimates. The former satisfy regularity conditions at all data points for

[10] There are 30 rather than 36 distinct models because (a) for the LES, translating and scaling yield identical models and (b) translating cannot be applied to the BTL without adding translating parameters and when translating parameters are introduced into the BTL it becomes the GTL.

[11] Because our estimates are based on aggregate rather than household data, there is no theoretical presumption that these demand functions correspond to well-behaved preferences.

[12] As shown in Chapter 5, the likelihood function for the maximum likelihood procedure contains two terms that do not appear with the generalized first difference procedure; therefore it is not clear a priori which likelihood function will be larger.

[13] We use the standard chi-square test to determine whether a restriction can be imposed.

Table 2a Log Likelihood Values: Generalized First Difference Procedure

	LES	QES	BTL	GTL
Static				
R = 0	266.72 (16)	276.11 (0)	282.96 (16)	286.96 (12)
R diagonal	299.38 (35)	302.99 (35)	307.75 (0)	309.48 (0)
R free	306.44 (35)	309.53 (35)	309.53 (35)	316.22 (0)
Dynamic translating				
R = 0	301.42 (20)	306.36 (0)	+	309.645 (14)
R diagonal	306.19 (35)	309.89 (35)	+	312.356 (11)
R free	310.27 (35)	315.19 (35)	+	*
Dynamic scaling				
R = 0	301.42 (20)	305.07 (33)	306.29 (23)	309.751 (10)
R diagonal	306.19 (35)	310.32 (35)	311.11 (22)	312.22 (16)
R free	310.27 (35)	314.82 (35)	314.34 (35)	316.75 (35)

Notes:
1. See Table 1.
2. The numbers in parentheses indicate at how many of the 35 sample price–expenditure situations preferences are quasiconcave.
3. The asterisk indicates failure to converge.

Table 2b Log Likelihood Values: Maximum Likelihood Procedure

	LES	QES	BTL	GTL
Static				
R = 0	272.24 (15)	281.60 (0)	289.33 (15)	292.82 (11)
R diagonal	291.94 (35)	299.69 (12)	300.99 (15)	302.24 (10)
R free	309.83 (35)	313.29 (35)	312.18 (35)	314.98 (35)
Dynamic translating				
R = 0	310.56 (20)	312.00 (26)	+	315.91 (17)
R diagonal	311.39 (33)	314.03 (0)	+	316.74 (17)
R free	312.49 (35)	317.04 (0)	+	*
Dynamic scaling				
R = 0	310.56 (20)	312.79 (32)	313.26 (23)	315.53 (11)
R diagonal	311.39 (33)	313.74 (35)	315.88 (23)	316.22 (15)
R free	312.49 (35)	318.48 (19)	319.17 (20)	320.31 (20)

Note:
1. See Table 2a.

both translating and scaling, while the latter satisfy them at no data points under translating and at 19 of 35 under scaling.

We conclude then that the QES with either dynamic translating or scaling and a free R matrix is our preferred model. The estimation procedure does not appear to affect this general conclusion.

The results for the BTL–GTL estimates are not as clear-cut as those for the LES–QES estimates. The first problem encountered is that neither estimation procedure converged for the GTL with R free under dynamic translating.[14] Thus we do not report results for this model under either estimation procedure. Our most general model then is the GTL with dynamics incorporated through scaling and with R free. As we will see shortly some of our conclusions regarding the translog form are sensitive to the estimation procedure, and/or to the order of testing. There is, however, one result that is robust wth respect to such considerations: in no case can the restrictions corresponding to the simple null hypothesis of a BTL form be rejected.[15] Thus in the following discussion we consider only the BTL results with dynamic scaling. The acceptability of further restrictions on our most general model—the BTL with R free under dynamic scaling—depends on the estimation procedure and, in one case, on the order of testing. For the ML estimates, the null hypothesis of no serial correlation cannot be rejected, after which the static model is rejected: thus, our preferred model incorporates dynamic scaling but no serial correlation of the residuals. For the generalized first difference method we can follow two different test routes. First we can accept R diagonal and then accept the static model, in which case our preferred model is static with a single correlation coefficient for the residuals. Alternatively we can accept the static model with R free but then reject R diagonal. Thus, using the generalized first difference estimation procedure leads us to conclude that the static model, with R diagonal or free, adequately reflects behavior.

The fact that the two estimation procedures yield different conclusions suggests that it is difficult with this functional form to disentangle the dynamic aspects appearing in the stochastic and nonstochastic parts of the model. In one case we have lagged consumption and no serial correlation of the disturbances, and in the other we have no lagged consumption and serial correlation of the disturbances (R either diagonal or free).

Having considered the LES–QES and BTL–GTL forms separately we now make a few comparisons between them. Because the LES is nested in the GTL, a standard likelihood ratio test of the LES restrictions is appropriate. These restrictions involve setting six parameters to 0, with a corresponding 5% chi-square critical value of 12.6. From Tables 2a and 2b it can be seen that in only 5 of the 18 comparisons is the GTL a significant improvement over the LES, and all of these involve static models. This, together with the fact that the GTL is never a significant improvement

[14] It should be recalled that the GTL contains three more parameters than the QES—in this case 18 as opposed to 15 in the most general QES models, and with the ML method three additional covariance parameters must be estimated.

[15] This finding differs from that in Chapter 6 where we found the GTL to be superior to the BTL. The difference probably lies in the greater degree of expenditure variation in the cross-section data used in Chapter 6.

over the BTL, suggests that, at least with this time series sample, the GTL is not a very useful form.

The likelihood ratio test cannot be used to compare the QES and the BTL because they are not nested. They do, however, involve the same number of independent parameters, so we can compare them informally on the basis of their likelihood values. This comparison reveals that in 10 of the 12 cases the BTL has a higher likelihood than the QES.

We consider now the effect of different estimation procedures and different functional forms on estimated marginal budget shares and own-price elasticities. In Table 3 we compare the QES with the BTL, each estimated by the generalized first difference and maximum likelihood methods. The models are estimated with dynamic scaling and R free. A comparison of estimation methods reveals substantial differences in both the pattern over time and in the magnitude of some of the estimates. For example, the marginal budget share for food generally declines over time with the maximum likelihood estimates, but generally increases over time with the first difference method. A priori one would expect this share to decline over time as real income rises. The own-price elasticity estimates for food are about the same with the two methods but those for clothing

Table 3 Marginal Budget Shares and Own-Price Elasticities

	Marginal budget shares			Own-price elasticities		
Year	Food	Clothing	Miscellaneous	Food	Clothing	Miscellaneous
			QES (maximum likelihood)			
1950	.52	.33	.15	−.62	−.40	.07
1960	.48	.27	.25	−.61	−.40	−.25
1970	.47	.25	.28	−.64	−.43	−.34
1980	.48	.23	.30	−.67	−.50	−.41
			QES (first difference)			
1950	.39	.24	.37	−.65	−1.02	−.66
1960	.39	.24	.37	−.65	−1.02	−.65
1970	.41	.20	.39	−.70	−.96	−.65
1980	.47	.13	.41	−.71	−.88	−.63
			BTL (maximum likelihood)			
1950	.52	.28	.20	−.67	−.56	−.04
1960	.50	.26	.24	−.68	−.53	−.23
1970	.47	.25	.27	−.67	−.52	−.35
1980	.47	.23	.30	−.67	−.48	−.43
			BTL (first difference)			
1950	.30	.16	.53	−.74	−.74	−.80
1960	.27	.16	.57	−.71	−.74	−.83
1970	.27	.18	.55	−.63	−.69	−.79
1980	.40	.14	.46	−.57	−.40	−.64

and miscellaneous are considerably higher with the first difference procedure.

A comparison of functional forms using the maximum likelihood procedure reveals estimates that are roughly the same. The pattern over time of the marginal budget shares is the same with the two forms and indeed the magnitudes are comparable. The own-price elasticities are remarkably similar, although they are generally slightly higher for the BTL.[16] A comparison of functional forms using the first difference method reveals substantial differences in both the magnitude and pattern over time of the marginal budget shares and in the magnitude of the own-price elasticity for clothing.

The treatment of the first observation may affect the ranking of functional forms and predicted behavior because some of the roots of the R matrix may be near unity for the generalized first difference method. If this is the case, then the term in the likelihood function that restricts the roots to lie in the unit circle under the maximum likelihood procedure may lead to convergence at quite a different parameter vector. To investigate this possibility, we have calculated the roots of the estimated R matrices for the four cases reported in Table 3. For the QES the largest roots are .67 and .98 for the maximum likelihood and generalized first difference methods, respectively, while for the BTL they are .68 and .99. These substantial differences in the roots of the R matrix with the two estimation methods are consistent with our finding that the two methods, differing only in their treatment of the first observation, can yield different predictions about behavior.

Although aggregate demand behavior need not satisfy regularity conditions even if the behavior of each household does, demand systems that violate regularity conditions may imply implausible responses to price or expenditure changes. Tabulating the number of sample price–expenditure situations at which regularity conditions are satisfied, we find that of a total of 420 possible situations the LES, QES, BTL, and GTL satisfy the conditions at 349, 306, 227, and 175 points, respectively.[17] On these grounds, then, the linear and quadratic expenditure models fare better than the translog models.

Finally we have investigated the dynamic stability of the nonstochastic models using simulations based on the estimated parameters. With prices and expenditure fixed at actual levels for the final year of the sample, all models converge to a steady state.

Although the generalized first difference and maximum likelihood procedures yield different estimates and different rankings of functional

[16] The QES maximum likelihood estimates predict positive own-price elasticities for the first four years of the sample.

[17] We exclude dynamic translating models from this comparison because they are not defined for the BTL.

forms in some cases, it is not clear which procedure is better. The maximum likelihood procedure requires strong assumptions about the stationarity of the autoregressive process and the normality of the disturbances. In particular, it assumes that the autoregressive process is stationary, whereas the generalized first difference procedure does not.[18] Hence, if the process generating the disturbances changed shortly before the sample period, then the maximum likelihood procedure involves a specification error, while the generalized first difference procedure does not. Since our sample period (1948–1983) begins shortly after World War II, it is plausible that the process generating the disturbances changed shortly before the beginning of our sample period.[19] The maximum likelihood procedure also assumes that the disturbances have a multivariate normal distribution. The generalized first difference procedure can be interpreted as an iterative Zellner procedure that iterates on the elements of the covariance matrix and R as well as on the structural parameters. Thus interpreted, the generalized first difference procedure does not require normality.

Since our sample size—35 price–expenditure situations—is small, even if we accept stationarity and normality we need not prefer maximum likelihood to generalized first difference estimates.[20] The desirable efficiency properties of maximum likelihood hold only asymptotically. Monte Carlo evidence would be useful, but we know of only one study dealing with equation systems. Maeshiro [1980] investigates the properties of various estimators for a nonsingular two-equation linear system assuming that the disturbances have a diagonal serial correlation matrix and that the contemporaneous covariance matrix is known. From his Monte Carlo analysis he concludes that with small samples and trended explanatory variables it is extremely important to use an estimator, such as maximum likelihood, that takes full account of the first observation. Further, of the estimators he studies, the one corresponding most closely to our generalized first difference procedure frequently performs worst.

Similarly, Monte Carlo evidence from single-equation linear models strongly suggests that when the independent variables are trending, retaining the first observation reduces substantially the root mean square

[18] By stationarity we mean that the process has been operating for a long time and that the characteristic roots of R [Chapter 5, Eq. (7)] lie within the unit circle. Without stationarity Eq. (8) of Chapter 5 does not hold and the first observation cannot be used separately in the estimation.

[19] Theil [1971, p. 253] and Poirier [1978] make similar arguments in the single-equation case.

[20] In the single-equation autoregressive case, Park and Mitchell [1979] report a Monte Carlo study using data generated from a stationary normal distribution; they find the iterative Prais–Winston procedure slightly superior to maximum likelihood.

error of parameter estimates.[21] To the extent that these results apply to nonlinear equation systems, maximum likelihood may be superior.

In summary we have investigated alternative functional forms, dynamic structures, and stochastic structures for demand systems. In addition we have investigated alternative estimation procedures differing in their treatment of the first observation. We find the QES to be superior to the LES and the GTL not superior to the BTL. Our preferred QES model incorporates either dynamic translating or scaling and a free R matrix; this conclusion is independent of the method of estimation. For the BTL our preferred model depends on the estimation method. For the maximum likelihood procedure it involves dynamic scaling but no serial correlation of the residuals. For the generalized first difference method it involves a static model with R either diagonal or free. The fact that the two estimation procedures yield different conclusions suggests that it is difficult with this functional form to disentangle the dynamic aspects appearing in the stochastic and deterministic parts of the model.

2. POOLING INTERNATIONAL CONSUMPTION DATA

In this section we report tests of pooling per capita time-series consumption data. Since pooling is most plausible for countries at similar stages of development, we use data from three advanced industrial societies: Belgium, the U.K., and the U.S. A nonparametric revealed preference test shows that the data from these three countries could not have arisen from maximizing a single, static, nonstochastic utility function. Thus we are led to incorporate dynamic elements and/or permanent differences among countries into the utility function. We estimate the QES using various specifications that permit both short-run and long-run differences among countries in an attempt to confirm the validity of pooling.

2.1. Demand System Differences

When countries' demand systems are identical, efficient estimation requires pooling. When countries' demand systems are unrelated, pooling is impossible: knowing one country's demand system parameters provides no information about any other's.[22] Between these extremes of identical

[21] See, for example, Beach and MacKinnon [1978], Harvey and McAvinchey [1978], Maeshiro [1979], and Park and Mitchell [1980]. In the single-equation case our generalized first difference procedure corresponds to the iterative Cochrane–Orcutt method. Although we have not presented results for the system analogue of the iterative Prais–Winston method, there is evidence (Park and Mitchell [1979]) that in the single-equation case it is very similar to maximum likelihood in small samples.

[22] We ignore the possibility of efficiency gains from taking account of nonzero disturbance covariances among countries.

and unrelated demand systems is a continuum of cases in which there are gains from pooling. We consider two polar cases permitting pooling and a comprehensive case that includes them both. Our two polar cases are the "permanent difference" specification in which some demand system parameters differ across countries and the "transitory difference" specification in which the underlying parameters are identical in all countries, but each country's short-run consumption pattern depends on its own past consumption.

Although our specifications of demand system differences can be applied to any demand system, we introduce them using the QES. The QES in share form is given by Eq. (5) in Chapter 6 and is not repeated here.

The permanent difference specification postulates that some subset of demand system parameters differs across countries, while the remaining parameters do not. For example, with the QES we might postulate that the b's differ across countries while the a's and c's are the same everywhere or, alternatively, that the a's and b's differ across countries while the c's are the same.[23] Since theoretical considerations do not dictate which parameters differ, the choice must be based on econometric convenience and empirical plausibility. Provided that there is some specified subset of demand system parameters that are identical across countries, efficient estimation requires pooling.[24]

The transitory difference specification postulates that some subset of demand system parameters depends on past consumption while the remaining parameters are constant and identical in all countries. For example, in the QES we might postulate that the b's depend linearly on consumption in the previous period:

$$b_{it} = b_i^* + \beta_i x_{it-1}. \tag{7}$$

The transitory difference specification assumes that the underlying parameters, the a's, b*'s, c's, and β's, are the same in all countries. This yields a model in which short-run demand responses to changes in prices and expenditure differ across countries because past consumption patterns differ, although the long-run or steady-state demand behavior is the same in all countries. The transitory difference specification implies a dynamic model in which the consumption pattern corresponding to a particular

[23] The parameters that differ need not constitute a "natural" subset: we might postulate that a_1, a_2, and b_1 differ, but that the c's and the remaining a's and b's are the same everywhere.

[24] The usefulness of a particular specification depends on the demand system functional form and the subset of parameters assumed to vary from one country to another. Suppose we have only one budget study from each of several countries and wish to estimate the LES. We can do so if we assume that all countries have identical demand systems, or, using the permanent difference specification, if we assume that the a's vary while the b's are the same in all countries. If we assume that the b's vary while the a's are the same, we cannot identify the b's. Thus, in conjunction with specific assumptions about which parameters vary from one country to another, the permanent difference specification may prove useful in pooling household budget data from different countries.

price–expenditure situation depends on past consumption. Since empirical studies using per capita time-series data from a single country have generally found dynamic specifications significantly superior to static ones, the transitory difference specification is attractive for pooling per capita time-series data.[25] Alternative transitory difference specifications correspond to different assumptions about which demand system parameters depend on past consumption and about the form of this dependence. Using the previous period's consumption to represent past consumption is convenient, but weighted averages of past consumption or the highest previous level are plausible and tractable alternatives.

The permanent difference specification is static—short-run and long-run demand functions coincide. It does, however, allow persistent differences across countries in demand functions, perhaps reflecting underlying differences in climate, in stable or slowly changing demographic characteristics, or in household production technology or tastes. The transitory difference specification is dynamic: countries with different consumption histories exhibit different consumption patterns when they face identical price–expenditure situations; in the long run, however, the consumption patterns of countries facing identical price–expenditure situations will converge. Although the transitory difference specification is compatible with habit formation or interdependent preferences, lagged consumption may be statistically significant in empirical demand analysis because of omitted variables unrelated to taste change; hence, although we find endogenous tastes a plausible hypothesis, we avoid equating the transitory difference specification with endogenous tastes.

The comprehensive specification combines the permanent and transitory difference specifications, allowing both persistent long-run differences in demand and a dynamic structure. For the QES one such comprehensive specification is obtained by modifying the transitory difference specification by assuming the b^*'s differ across countries, while still requiring the a's, c's, and β's to be the same in all countries. Alternative comprehensive specifications are obtained by assuming, for example, that only the a's differ across countries, or that only the c's differ, or that both the a's and c's differ.

2.2. Data and Results

2.2.1. *Preliminaries*

In this section we report tests of pooling per capita consumption data from Belgium, the U.K., and the U.S. for the years 1961–1978. We consider

[25] Our transitory difference specification is similar to the Anderson and Blundell [1982] approach (i.e., embedding steady-state solutions in a more general dynamic framework). The transitory difference specification is unlikely to be useful for analyzing household budget data because such data sets seldom report past consumption.

three commodity groups: food, clothing, and a nondurable miscellaneous group that excludes housing and educational expenditures.

Data availability determined our choice of countries. First, we considered only the 16 countries for which Kravis, Heston, and Summers [1978] report purchasing power parity (PPP) data.[26] Second, since pooling is most plausible for countries at similar stages of development, we considered only advanced industrial societies. Third, since pooling requires comparable commodity groups, we considered only countries for which the OECD National Accounts 1961–1978 [1980] reports data from which we could construct comparable food, clothing, and miscellaneous categories. Finally, since we wished to test the appropriateness of pooling, we required a sufficiently long time series for each country to estimate its demand system separately; although consumption data for several countries were available for the period 1970–1978, we considered only countries for which data were available for the full 18-year period.[27]

The demand system we estimate, the QES, is given (in share form) by Eq. (5) in Chapter 6. We obtain a stochastic specification by adding a disturbance term to each share equation. Denoting the $n \times 1$ vector of disturbances $(u_{t1}, \ldots, u_{tn})'$ by $\tilde{u}_t$, we assume that the $\tilde{u}_t$ are independently distributed across countries but that within each country they are correlated over time according to the simple first order scheme

$$\tilde{u}_t = \tilde{R}\tilde{u}_{t-1} + \tilde{e}_t, \tag{8}$$

where the $\tilde{e}_t$ are independently normally distributed with covariance matrix $\tilde{\Omega}$ and $\tilde{R}$ is a diagonal matrix with ρ on the main diagonal.[28] We obtain maximum likelihood estimates of the demand system parameters (with b related to past consumption as in (7)) of Ω and of ρ by maximizing the concentrated likelihood function associated with the model using all observations, including the first.[29]

2.2.2. *Pooling Three Countries*

Using the nonparametric revealed preference test procedure of Afriat [1967], Diewert [1973], and Varian [1982], we find that the data for each country separately could be generated by maximizing a static, nonstochastic utility function, but that the pooled data for all three countries could not be generated by maximizing a single, static, nonstochastic utility function.[30] Thus pooling data from different countries, if it is possible at

[26] As an alternative to using published PPP data, in Appendix C we describe a procedure that could be used in some cases to estimate PPPs along with demand system parameters.

[27] A complete description of data sources appears in Appendix B.

[28] We also allowed ρ to differ across countries, but since this had very little effect on the likelihood values, we have not reported those results here.

[29] As in Chapter 5, Ω corresponds to the first $n - 1$ rows and columns of $\tilde{\Omega}$.

[30] We are grateful to Hal Varian for making his programs available to us.

all, requires a specification that permits demand system parameters to differ across countries or to change over time.

An alternative approach to determining whether data from the three countries can be pooled is to estimate a utility function of a particular functional form. We estimate the QES for each country separately and for all three countries combined, allowing neither transitory nor permanent differences among countries, and use a standard likelihood ratio test to determine if the data from all three countries are generated by the same process. Pooling three countries constrains 8 demand system parameters (3 b's, 3 c's, and 2 a's) and 4 covariance parameters (3 elements of Ω and the value of ρ) to be the same in all countries; thus the likelihood ratio test of equality for three countries uses 24 degrees of freedom (two pairwise country comparisons involving 12 parameters for each country). Since our calculated chi-square value is 115.9 while the critical 1% value with 24 degrees of freedom is 43.0, we reject pooling of all three countries, a result consistent with that obtained using Varian's nonparametric test.[31,32] Thus we are led to try a specification that allows for transitory and/or permanent differences among countries.

We allow for transitory differences among countries by estimating the QES with dynamic translating. More specifically, as in Eq. (7) we assume that the b's for each good are linearly related to consumption of that good in the previous period, a specification we have called "linear dynamic translating" in Chapter 4.[33] With this dynamic specification the demand system for each country contains 4n independent parameters, including a serial correlation coefficient. We use the standard likelihood ratio test to determine whether the demand system parameters are the same in all three countries. Our calculated chi-square value is 119.0 while the critical value is 50.9. Thus we reject pooling data for all three countries

[31] We use the 1% critical values in an attempt to offset the small sample bias of the likelihood ratio test. Using the LES in a Monte Carlo experiment, Wales [1984] tested the (true) null hypothesis of no difference in parameter values for two data sets of size 20 at the 1% level and rejected the null hypothesis about 4% of the time. Thus our best guess is that using the 1% critical values corresponds to a Type I error of about 4%.

[32] Although the likelihood ratio test rejects pooling of all three countries using a static specification without permanent differences, the estimated pooled static QES does satisfy regularity conditions at all price–expenditure situations in the sample. This result does not contradict the nonparametric test result, but it does illustrate an important difference between the nonparametric and estimation approaches. The nonparametric approach examines regularity conditions using observed prices and observed quantities, while the demand system estimation approach uses observed prices and predicted quantities. That is, the estimation approach examines whether the estimated demand system is consistent with regularity conditions at each observed price–expenditure situation. The stringency of this procedure depends on the particular functional form: for example, if in the data all goods have nonnegative shares, then an estimated demand system corresponding to a Cobb–Douglas utility function must satisfy regularity conditions.

[33] We also estimated our model using the "dynamic scaling" procedure discussed in Chapter 4, but the results were so similar to those of dynamic translating that we have not reported them here.

using the QES with dynamic translating. The nonparametric approach is applicable only to static specifications. It cannot be used to determine whether the pooled data for all three countries could have been generated by maximizing a single dynamic utility function with transitory differences introduced through dynamic translating. Thus, continuing with our parametric approach, we now turn to a specification that allows for permanent differences among countries.

Comprehensive specifications allow both a dynamic structure within countries and long-run differences among countries. We next consider four comprehensive specifications, each of which allows one subset of parameters (either the a's, b's, c's, or β's) to vary among countries but assumes that the remaining parameters are identical in all three countries.[34] We see no a priori reason to prefer one of these specifications to the others. Letting only the b's, c's, or β's vary involves 24 restrictions and a critical chi-square value of 43.0, while our calculated values are 52.9, 82.2, and 67.8, respectively. Letting the a's vary involves 26 restrictions and a critical value of 45.6, while our calculated value is 67.0. Thus, even using the comprehensive specifications, the likelihood ratio test rejects the pooling of all three countries.

2.2.3. Pooling Two Countries

The results for pooling pairs of countries are more complex. Nonparametric tests show that the data from one pair of countries (U.K.–U.S.) are consistent with maximization of a static, nonstochastic utility function, while data from the other two pairs of countries (Belgium–U.K., Belgium–U.S.) are not. Estimating the QES for each of the three pairs of countries without allowing either permanent or transitory differences, we find that the likelihood ratio test rejects pooling for all three pairs. This result appears in the first row of Table 4, which records the calculated chi-square values used in the likelihood ratio tests. Although pooling is rejected for the U.K.–U.S. pair, the rejection is far less decisive than it is for the other two pairs of countries.

Since the nonparametric test shows that the U.K.–U.S. data are consistent with maximizing some utility function, the rejection of pooling with the QES could be interpreted as evidence of functional form misspecification. The rejection, however, might also be interpreted as a reflection of dynamic misspecification. Suppose, for example, that the data were generated by a dynamic QES; for a particular set of price–expenditure

[34] The only other reference in the literature of which we are aware that involves this type of testing is Goldberger and Gamaletsos [1970]. Discussing the possibility of pooling data for 13 OECD countries, these authors conclude "computations not reported here indicate rejection of the hypothesis that all parameters are fixed across countries, but leave open the possibility that subsets of parameters possess that invariance property" (p. 381).

Table 4 Chi-Square Values for Pooling Pairs of Countries

	DOF	B.–U.K.	U.K.–U.S.	B.–U.S.
Static utility function	12	82.1	35.2	65.6
Dynamic utility function				
No parameters differing	15	76.1	77.4	68.0
b's differing	12	35.2	23.9	23.9
c's differing	12	43.5	36.2	45.5
β's differing	12	40.2	52.9	29.1
a's differing	13	35.6	44.8	38.4

Note:
1. DOF denotes degrees of freedom for the likelihood ratio test. The 1% critical values for 12, 13, and 15 degrees of freedom are 26.7, 27.7, and 30.6, respectively.

situations, the price–quantity data might pass the nonparametric consistency test while failing the likelihood ratio test for pooling based on the estimated but misspecified static QES. Alternatively, the rejection might be interpreted as a reflection of a failure to specify appropriately permanent differences among countries. Although one's first instinct is that Occam's razor argues for treating constant tastes as a maintained hypothesis, this is not the case. Such an argument depends on the implicit assumption that constant tastes is the simpler hypothesis—regardless of the complexity of the static utility function required to rationalize observed behavior.

As in the case of pooling three countries we next consider a specification that allows transitory but not permanent differences between countries. We estimate the QES with dynamic translating for each country separately and for each pair of countries and use likelihood ratio tests to assess the validity of pooling. For each of these tests, the calculated chi-square value, which appears in the second row of Table 4, is over twice as large as the corresponding critical value. Thus we reject pooling pairs of countries for the QES with dynamic translating and conclude that, with our data and demand system, allowance must be made for permanent differences between countries.

As in the three-country case, our comprehensive specifications allow for both a dynamic structure and long-run differences between countries. The calculated chi-square values for the four comprehensive specifications, which appear in rows 3 through 6 of Table 4, indicate that we reject pooling in 10 of 12 cases. In particular, we do not reject pooling for the Belgium–U.S. and the U.K.–U.S. pairwise comparisons when the b's are allowed to differ. We reject pooling in all other cases. That is, we reject pooling when parameter sets other than the b's are allowed to vary, and we reject pooling of the Belgium–U.K. pair when the b's are allowed to differ.

To predict elasticities and marginal budget shares we use the pooled Belgium–U.S. estimates for Belgium and the pooled U.K.–U.S. estimates

Table 5 Marginal Budget Shares and Own-Price Elasticities: Pooled Data

Marginal budget shares			Own-price elasticities		
Food	Clothing	Miscellaneous	Food	Clothing	Miscellaneous
Belgium (based on pooled Belgium–U.S. data)					
.38	.18	.44	−.42	−.28	−.54
.35	.18	.47	−.47	−.45	−.67
.31	.18	.51	−.50	−.60	−.78
United Kingdom (based on pooled U.K.–U.S. data)					
.15	.24	.61	−.10	−.54	−.78
.13	.21	.66	+.09	−.46	−.65
.17	.18	.65	−.07	−.45	−.65
United States (based on pooled U.K.–U.S. data)					
.26	.19	.55	−.79	−.81	−1.05
.23	.19	.57	−.68	−.72	−.95
.27	.17	.57	−.66	−.67	−.88
United States (based on pooled Belgium–U.S. data)					
.25	.18	.57	−.64	−.71	−.92
.21	.18	.61	−.60	−.73	−.93
.18	.17	.65	−.56	−.75	−.93

Notes:
1. Demand system parameters were estimated using pooled data from two countries; marginal budget shares and elasticities were evaluated at the price–expenditure situation of the designated country for the first, middle, and last year of the sample.
2. The demand system estimated using pooled Belgium–U.S. data satisfies regularity conditions at all 17 Belgian price–expenditure situations in our sample. Similarly, the estimates using pooled U.K.–U.S. data satisfy regularity conditions at all 17 U.S. sample points, and the estimates using Belgium–U.S. data satisfy regularity conditions at all 17 U.S. sample points. The demand system estimated using pooled U.K.–U.S. data, on the other hand, violates regularity conditions at all 17 U.K. price–expenditure situations in our sample.

for the U.K., with the b's allowed to vary. These estimates are presumably more efficient than those based on individual country data alone.[35] For the U.S., with two conflicting sets of pooled estimates, there is no conventional solution, and we present estimates based on both. In Table 5 we report estimated marginal budget shares and own-price elasticities at three price–expenditure situations for each country, those corresponding to the first, middle, and last years of our sample. All estimated marginal budget shares and elasticities appear reasonable a priori except for the positive own-price elasticity for food in the U.K. in the middle sample period.[36] For the U.S. the two sets of estimates are very similar at the middle sample point, but in several cases the trends over the sample period differ.

[35]When countries are treated separately we estimate 15 parameters using 18 observations (i.e., 18 years of price and quantity data on three goods), while when countries are treated in pairs with the b's varying we estimate 18 parameters using 36 observations (on three goods).

[36]Although not shown in the table, this positive elasticity does not differ significantly from 0 at the 5% level.

Table 5 also reports the number of sample price–expenditure situations at which the estimated demand system parameters correspond to well-behaved preferences. For the U.S. and for Belgium, these regularity conditions are satisfied at all 17 sample points for both sets of estimates; for the U.K. they are not satisfied at any sample points.[37]

Finally we report on our tests of functional form and dynamic structure. We find the QES functional form significantly superior to the LES, the null hypothesis being that $c_1 = c_2 = c_3 = 0$; the calculated chi-square values for the Belgium–U.S. and U.K.–U.S. pairs are 29.3 and 19.9, while the critical 1% value is 6.6, thus indicating significantly nonlinear expenditure responses. For dynamic structure our findings are mixed. To test whether the short-run and long-run demand functions differ, the null hypothesis is $\beta_1 = \beta_2 = \beta_3 = 0$. The calculated chi-square values for the Belgium–U.S. and U.K.–U.S. pairs are 2.6 and 22.5, while the 1% critical value is 6.6. Thus, using the pooled estimates, we reject the hypothesis that short-run and long-run demands are the same for the U.K., but we do not reject it for Belgium; for the U.S. we have conflicting results.

In summary we find that even under circumstances favorable to pooling—using countries at the same general stage of development—pooling is not generally acceptable. Nonparametric revealed preference tests show that the pooled data from Belgium, the U.K., and the U.S. could not have been generated by maximizing a single, static, nonstochastic utility function. Our attempts to pool data using a parametric approach based on the QES, allowing for both permanent differences among countries and a dynamic structure within countries, were generally unsuccessful.

We have considered models in which no parameters vary across countries and those in which only one set varies; for most parameter sets and combinations of countries, we rejected pooling. We find this result disappointing since our specifications were designed to meet the legitimate objection that it is implausible to assume that all countries have identical short-run and long-run demand behavior. Furthermore, the three countries we analyzed represented a case favorable to pooling. There is no reason to believe we would have done better pooling data from Canada, Germany, and the Netherlands, for example, and some reason to believe we would have done worse pooling data from France, India, and Kenya.[38]

[37]We calculate regularity conditions at 17 rather than 18 sample points because, in our dynamic models, calculating regularity conditions requires the previous period's consumption pattern.

[38]Our modeling of demand differences between countries may be too simple; it is possible that we would do better allowing more than one parameter set to vary. Thus, we could allow two parameter sets to vary across countries, say the b's and c's, while requiring the β's and a's to be the same. We have estimated specifications of this type for all six pairs of parameter sets and all three pairs of countries. Although detailed results are not reported here, we find that pooling is permissible in 11 of the 18 cases. Pooling data for all three countries is, however, never permissible.

Our results suggest that researchers should be cautious about pooling data from different countries for demand system estimation.

APPENDIX A: U.S. DATA, 1948–1983

Constant (1972) and current dollar expenditures on the various categories of goods for the period 1948–1975 were obtained from Tables 2.7 and 2.6, respectively, of *The Survey of Current Business*, July 1978. In terms of the categories, we defined our three commodity groups as follows (numbers in parentheses correspond to those in Table 2.7):

A. Food
   1. Food (17)

B. Clothing
   1. Clothing and shoes (23)
   2. Shoe cleaning and repair (61)
   3. Cleaning, laundering, dyeing, etc. (62)

C. Miscellaneous
   1. Toilet articles and preparations (31)
   2. Tobacco products (30)
   3. Drug preparations and sundries (34)
   4. Nondurable toys and sports supplies (35)
   5. Domestic service (48)
   6. Barbershops, beauty parlors, and baths (63)
   7. Medical care services (64)
   8. Admissions to specified spectator amusements (69)

Constant (1972) and current dollar expenditures on the various categories of goods for the period 1976–1983 were obtained from Tables 2.4 and 2.5, respectively, of *The Survey of Current Business*, July 1984. In terms of the categories, we defined our three groups as follows (numbers in parentheses correspond to those in Table 2.5):

A. Food
   1. Food (20)

B. Clothing
   1. Clothing and shoes (27)
   2. Cleaning, storage, and repair of clothing and shoes (68)

C. Miscellaneous
   1. Toilet articles and preparations (35)
   2. Tobacco products (34)
   3. Drug preparations and sundries (38)
   4. Nondurable toys and sports supplies (39)
   5. Domestic service (54)
   6. Barbershops, beauty parlors, and baths (69)

7. Medical care services (71)
8. Admissions to specified spectator amusements (83)

Per capita consumption of each good was calculated by dividing annual expenditure in 1972 dollars by population. The population figures are "resident population" in the U.S. and are taken from Table 2 of *The Statistical Abstract of the U.S., 1976*, for the period 1948–1972 and from *Current Population Reports*, Series P 25, No. 983, June 1984 for the period 1973–1983. Price indexes were determined by dividing current dollar expenditure by constant dollar expenditure for each of the three categories.

APPENDIX B: OECD DATA, 1961–1978

Consumer Expenditure Data

Expenditure in current prices and constant prices on the various categories of goods were obtained from Tables 4a and 5b of the *OECD National Accounts 1961–1978*, Volume 2. In terms of the categories, we defined our three broad commodity groups as follows (numbers in parentheses correspond to those in Tables 5a and 5b):

A. Food
   1. Food (1-1)
   2. Nonalcoholic beverages (1-2)
   3. Alcoholic beverages (1-3)
B. Clothing
   1. Clothing and footwear (2)
C. Miscellaneous
   1. Other recreation, etc. (7-2)
   2. Miscellaneous goods and services (8)
   3. Tobacco (1-4)

Per capita expenditure on each commodity group was calculated by dividing annual expenditure by mid-year population. The population data for each country are from the *U.N. Demographic Yearbook, 1979*. Price indexes were obtained by dividing current price expenditure by constant price expenditure; we normalized the price index series to unity in 1970 by dividing each series by its 1970 value.

Purchasing Power Parity Data

The PPP data are from Tables 5.1 and 5.15 of Kravis, Heston, and Summers [1978]. In terms of the categories our three groups were defined as follows:

A. Food
   1. Food

2. Beverages

B. Clothing

1. Clothing

C. Miscellaneous

1. Recreation
2. Other expenditures (personal care and miscellaneous services)
3. Tobacco

The resulting categories appear very similar to those constructed from the expenditure data. We aggregated these PPP data to our commodity group levels by combining the country-weighted PPPs from the tables using the corresponding country expenditures as weights to give country-weighted PPPs at the commodity group level. The PPPs we used are the geometric means of these country-weighted PPPs for each commodity group. To obtain comparable quantities, we divided the normalized Belgian and U.K. quantity data for food, clothing, and miscellaneous by the corresponding PPPs; comparable price series were obtain by multiplying the normalized price series by the PPPs.

APPENDIX C: ESTIMATING PURCHASING POWER PARITIES

When reported PPPs are not available, it is sometimes possible to estimate the required transformation factors along with the demand system parameters.[39] The PPPs—one for each good in each country, except the base country for which they are unity—enter the demand system by multiplying the prices and dividing the quantities.

We attempted to estimate PPPs for the U.K.–U.S. and Belgium–U.S. pairs analyzed above, using the QES in which the b's differ between countries. Because the PPPs and the b's enter the QES demand equations in very similar ways, separating their effects is bound to be difficult. We found that the likelihood function was extremely flat with respect to the PPPs and b's, and we were unable to converge the model. Indeed if the LES rather than the QES were being estimated the model would not be identified. This suggests that a specification allowing parameters other than the b's to vary might perform better. Since our data do not accept pooling when other single-parameter sets are allowed to differ, we have estimated specifications in which two sets of parameters are allowed to differ. For example, when we allow the a's and c's to vary, then pooling the U.S. and U.K. data is acceptable; with this model we have estimated PPP values for the U.K. of .46, .93, and .32 for food, clothing, and miscellaneous, respectively. Kravis, Heston, and Summers report values of .35, .33, and

[39]This technique for estimating unobserved transformation factors is analogous to demographic scaling.

.35, while the exchange rate in 1970 was .4174. Although the estimates for food and miscellaneous are fairly close to the reported values, the corresponding standard errors are .42, 1.6, and 3.9, so the point estimates are of limited interest. Nevertheless this illustrates a technique that might allow the pooling of other data sets in the absence of PPP information.

This page intentionally left blank