JOURNAL OF APPLIED ECONOMETRICS
J. Appl. Econ. **16**: 59–80 (2001)

AN EMPIRICAL COMPARISON OF FLEXIBLE DEMAND SYSTEM FUNCTIONAL FORMS

DOUGLAS FISHER,[a] ADRIAN R. FLEISSIG[b]* AND APOSTOLOS SERLETIS[c]

[a] *Department of Economics, North Carolina State University, Raleigh, North Carolina 27695-7506, USA*
[b] *Department of Economics, California State University, Fullerton, CA 92834-6848, USA*
[c] *Department of Economics, The University of Calgary, Calgary, Alberta T2N 1N4, Canada*

SUMMARY

This paper compares the performance of eight frequently used flexible forms that are either (1) locally flexible, (2) 'effectively globally regular', or (3) asymptotically globally flexible. Results show that the functions with global properties generally perform better, particularly those models having asymptotic properties. Results, using US consumption data, indicate substitutability among the components of consumption at most data points. There is also some interesting substitution volatility around the time of recessions in the USA.

1. INTRODUCTION

This paper presents an empirical comparison and evaluation of the effectiveness of some well-known flexible functional forms on US aggregate consumption data. There are two main purposes to our investigations. First and foremost, we want to investigate the differences among eight flexible functional forms as they perform on an interesting data set (the quarterly US aggregate consumption data for the period 1960:1 to 1991:4). Second, we are interested in what the better-performing flexible forms have to say about aggregate consumption in the United States. That is, while our first task provides the main focus of the paper, we try not to lose sight of the fact that this activity has a purpose, which is to find the best possible way to model US consumption behaviour.

The flexible functional forms examined are either parametric or semi-non-parametric. We want to emphasize that while the semi-non-parametric forms generally have more desirable asymptotic properties than the parametric functions, it does not follow that one is to be preferred to the other on particular data sets (such as the US aggregate consumption data). To effect such a comparison, we propose a variety of criteria, drawing on the literature of course, and using data that have been pre-filtered to be consistent with a well-behaved aggregate utility function over much of the sample.[1] The empirical problem, then, is to use the different functional forms to approximate the underlying indirect utility function for these data. The relative satisfaction of the regularity conditions for the aggregate consumer is what we will seek, although we will also look at expenditure responses over a range of expenditures, study the results of forecasting exercises, and examine the estimates of the elasticities of substitution to see if they are consistent with economic theory.

* Correspondence to: Adrian R. Fleissig, Department of Economics, California State University, Fullerton, CA 92834-6848, USA. E-mail: afleissig@fullerton.edu
Contract/grant sponsor: Social Sciences and Humanities Research Council of Canada.

[1] We are referring here to the use of the NONPAR procedure of Varian (1982, 1983), itself an application of the Generalized Axiom of Revealed Preference.

Received 20 November 1996
Revised 12 May 2000

The plan of this paper is on follows. In Section 2 we briefly sketch the theory behind our approach; our purpose here is to explain how our choices of flexible functional form fit into the theoretical literature and to try to explain why one form might be expected to perform better than another. Section 3 explains the specific characteristics of each of the eight flexible functional forms. The data set, and how it was constructed, is described in Section 4. The empirical part of the paper begins, in Section 5, with a comparison of the performance of the eight flexible functional forms on the quarterly US data, elasticities of substitution (in Section 6), and out-of-sample forecasts (in Section 7). Our brief conclusions appear in Section 8.

2. THE DEMAND SYSTEMS APPROACH

The demand-systems approach provides an effective method to impose and test neoclassical restrictions on individual behaviour; specifically the monotonicity and curvature restrictions.[2] A functional form is selected to approximate the indirect utility or cost function and then the corresponding demand or share equations are derived using Roy's identity or Shephard's lemma. There are, however, many functional forms that can be used. These flexible functional forms differ in their specific parameterization and approximation properties which are now briefly discussed.

Among the most popular of the earliest *locally flexible* functional forms are the generalized Leontief, translog, and Almost Ideal Model (AIDS) specifications. These locally flexible functional forms initially showed some promise but they have some troublesome limitations. For example, Caves and Christensen (1980), Barnett and Lee (1985) and Barnett *et al*. (1985) show that the regularity regions of local flexible functional forms can be relatively small. Furthermore, the Monte Carlo analyses of Guilkey and Lovell (1980) and Guilkey *et al*. (1983) find that the generalized Leontief and the translog fail to provide a satisfactory approximation to the true data-generating process for the moderate and even large elasticities of substitution that often arise in applications. Another troublesome result is that the translog can classify goods as complements when they are actually substitutes. Finally, an important reason for the failure of these locally flexible forms is that they can only provide a local approximation to the true data-generating function at a single point in a delta neighborhood of an unknown and often small size.

These problems led to the development of locally flexible functional forms that have larger regularity regions and higher rank models that can better approximate non-linear Engel curves. Cooper and McLaren (1996) discuss functions that have larger regularity regions that include all data points in the domain, as well as real expenditures, calculated from any combination of prices and nominal expenditures, exceeding the minimum value in the sample. Examples of these functions include the Laurent models introduced by Barnett (1983, 1985), Barnett and Lee (1985) and Barnett *et al*. (1985, 1987) and the General Exponential Form (GEF) of Cooper and McLaren (1996). The rank of a demand system, as discussed in Lewbel (1987a,b,1990,1991), has implications for aggregation and the non-linearity of Engel curves. Higher rank models, such as the Quadratic Almost Ideal Demand System (QUAIDS) of Banks *et al*. (1997) which is rank 3, can approximate more non-linear Engel curves often found in empirical analysis. They note that at sufficiently high expenditure levels, a QUAIDS budget share may violate the zero-to-one range. Nonetheless, it appears as though the regular region is considerably larger than the locally

[2] The monotonicity restriction requires that the values of fitted demand be non-negative. The curvature condition requires quasi-convexity of the indirect utility function.

flexible forms and thus we classify the QUAIDS model as effectively globally regular. While these models provide a better approximation over the initial flexible forms, they may not be asymptotically regular and may fail to provide an effective approximation of the derivatives—and hence the curvature of the true data-generating function.

Semi-non-parametric (SNP) functions can provide an asymptotically global approximation for complex economic relationships.[3] These SNP functions provide global approximations to the true data generating process and its partial derivatives. By *global approximation* we mean that the flexible functional form is capable, in the limit, of approximating the unknown underlying generating function at all points and thus of producing arbitrarily accurate elasticities at all points. Two such SNP functions are the Fourier flexible functional form (FFF) and the Asymptotically Ideal Model (AIM).

3. EIGHT FLEXIBLE FUNCTIONAL FORMS COMPARED

We now provide a theoretical comparison of the eight different functional forms just mentioned by grouping them into three sets that have broadly similar characteristics. These sets are (1) locally flexible forms, (2) effectively globally regular forms, and (3) asymptotically globally flexible. We selected these eight forms, even though there are many other possibilities,[4] because they provide a representation of the three groups of functional forms that are in the widest use in such studies.

3.1. Locally Flexible Functional Forms

The Generalized Leontief (GL) Model

The GL function, due to Diewert (1971), can be written as follows:

$$h(v) = \alpha + \sum_{i=1}^{n} \alpha_i v_i^{1/2} + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} v_i^{1/2} v_j^{1/2} \tag{1}$$

where v_i is the expenditure normalized price for good i. A sufficient condition for global regularity is that $\beta_{ij} > 0$, $\alpha_i > 0$ for all i and j. Caves and Christensen (1980) have shown that the GL has satisfactory local properties when preferences are nearly homothetic and substitution is low, implying that the GL can approximate Leontief preferences well. However, when preferences are not homothetic and substitution increases, they show that the GL has a rather small regularity region. Symmetry ($\beta_{ij} = \beta_{ji}$) and adding up ($\sum \alpha_i = 1$) restrictions (for $i = 1, \ldots, n$) are imposed in estimation.

The Basic Translog (BTL) Model

The BTL introduced by Christensen *et al*. (1975) approximates the reciprocal of the indirect utility function using a second-order Taylor series expansion:

$$\ln h(v) = \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln v_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} \ln v_i \ln v_j \tag{2}$$

[3] Elbadawi *et al*. (1983) define a semi-non-parametric function as a truncated series expansion that is dense in a Sobolev norm.

[4] Lewbel (1987b, 1990, 1991, 1995) classifies functional forms in terms of fractional demand systems, full rank demand systems, and Engel curve approximation. He speculates that the asymptotically global models should obtain asymptotic consistency while still maintaining integrability.

Symmetry, adding up, and homogeneity require that $\beta_{ij} = \beta_{ji}$ and $\sum \alpha_i = 1$ restrictions (for $i = 1, \ldots, n$) are imposed in estimation.[5] Guilkey *et al.* (1983) show that the translog is globally regular if and only if preferences are Cobb–Douglas, meaning that the translog performs well if substitution between all commodities is close to unity. They show that the regularity properties deteriorate rapidly when substitution diverges from unity.

The Almost Ideal Demand System (AIDS)

The AIDS model of Deaton and Muellbauer (1980) is a widely used flexible demand specification obtained from the following PIGLOG (price-independent generalized logarithmic) expenditure function:

$$\ln C(U, \boldsymbol{p}) = \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln p_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij} \ln p_i \ln p_j + U^* \beta_0 \prod_{k=1}^{n} p_k^{\beta_k} \tag{3}$$

where C is the minimum level of expenditure that is necessary to achieve utility level U^* at given prices. The demand equations in budget share form appear as follows:

$$s_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i \ln\left(\frac{m}{p}\right) \qquad (i = 1, \ldots, n) \tag{4}$$

Here P is a translog price index defined by

$$\ln P = \alpha_0 + \sum_{k=1}^{n} \alpha_k \ln p_k + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} \gamma_{jk} \ln p_j \ln p_k$$

Adding up, symmetry, and homogeneity restrictions require that $\sum \alpha_i = 1$, $\sum \gamma_{ij} = 0$, $\sum \beta_i = 0$, and $\gamma_{ij} = \gamma_{ji}$ for $i, j = 1, \ldots, n$. Since the estimation of the AIDS model is difficult using the translog price index, Stone's price index P^* is often used instead of P, where

$$\ln P^* = \sum_{k=1}^{n} w_k \ln p_k \qquad w_k = \frac{p_k X_k}{m}$$

are budget shares. Results from Pashardes (1993), Buse (1994), and Alston *et al.* (1994) show that using the Stone index approximation can severely bias the results. Therefore, our empirical work uses a procedure recommended by Pashardes (1993),[6] and does not use the Stone index approximation. However, even in this case, the approximation performance of the AIDS model may still be poor because it is a locally flexible form and may have a relatively small regularity region.[7]

[5] The BTL is a special case of the Generalized Translog (GTL) due to Pollak and Wales (1980). See Serletis (1988) for a comparison of various translog flexible forms.

[6] We multiply the translog price index by $\ln(p_{it})$, add α_0 to both sides, and then sum over i to get the following expression:

$$\ln P = \left(\alpha_0 - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij} \ln p_i \ln p_j - \ln m \sum_{i=1}^{n} B_i \ln p_i\right) \Bigg/ \left(1 - \sum_{i=1}^{n} B_i \ln p_i\right)$$

We then set α_0 equal to the minimum of $\ln(m)$. The AIDS model is then estimated, conditioned on the value for α_0, using the Stone index with the parameter estimates as starting values. This iterative procedure is repeated until the AIDS model converges; this is assured since a quadratic equation is minimized at each step.

[7] Ramajo (1994), using the Laurent series expansion, and Chalfant (1987), using Fourier series, show how to increase the regularity region of the AIDS model considerably.

As argued earlier, models such as the GL, BTL, and AIDS are locally flexible but may have a relatively small regular region.

3.2. Effectively Globally Regular Functional Forms

A partial solution to the problem just discussed has recently been provided by Barnett (1983), Banks *et al*. (1997), Cooper and McLaren (1996), and others. These authors, as discussed above, developed locally flexible functional forms with larger theoretical regularity regions that are capable of approximating more general Engel curves. These functions are labeled by Cooper and McLaren as effectively globally regular".

The Full Laurent Model

Barnett (1983), Barnett and Lee (1985), and Barnett, Lee, and Wolfe (1985, 1987) developed functional forms that employ the Laurent series expansion as the approximating mechanism. The full Laurent model,[8] defined by Barnett (1983), is based on the (second-order) Laurent reciprocal indirect utility function of

$$h(v) = a_0 + 2\sum_{i=1}^{n}(\alpha_i w_i - b_i w_i^{-1}) + \sum_{i=1}^{n}\sum_{j=1}^{n}(\alpha_{ij} w_i w_j - b_{ij} w_i^{-1} w_j^{-1}) \tag{5}$$

where $w_i = v_i^{1/2}$. The share equations (see Barnett, 1983) are homogeneous of degree zero in the parameters and require an arbitrary normalization.

Quadratic AIDS (QUAIDS)

Since Engel curves for US consumption data appear to be more non-linear than the rank two AIDS and translog models, Banks *et al*. (1997) develop a rank three demand system extension of the AIDS model, the Quadratic AIDS model. The indirect utility function for the QUAIDS model is

$$\log V(p, m) = \left(\frac{b(p)}{\log \dfrac{m}{a(p)}} - \lambda(p) \right)^{-1} \tag{6}$$

where $a(p) = \alpha_0 + \sum \alpha_a \log(p_a) + \sum\sum \gamma_{ar} \log(p_a)\log(p_r)$, $\log(b(p)) = \sum \beta_a \log(p_a)$, and $\log(\lambda(p)) = \sum \lambda_g \log(p_g)$. The corresponding share equations are:

$$w_s = \alpha_s + \sum_r \gamma_{sr} \log p_r + \beta_s \log \frac{m}{a(p)} + \gamma_s \frac{\left(\log \dfrac{m}{a(p)}\right)^2}{b(p)} \tag{7}$$

The constraints $\{\sum_s \alpha_s = 1, \sum_s B_s = 0, \sum_s \gamma_{sr} = 0, \sum_r \gamma_{sr} = 0, \sum_s \gamma_s = 0, \}$ are imposed so that the estimated demands satisfy the budget constraint and are homogeneous of degree 0 in prices and total expenditure.

The General Exponential Form (GEF)

Another flexible form that increases the range of Engel curves responses is the General Exponential Form (GEF) of Cooper and McLaren (1996) that has the utility function of $U(c, P) = (c - \kappa P1)/P2$. The price indices $P1$ and $P2$ are defined as CES:

$$Pk(p) = \left\lfloor \sum \beta k_i p_i^{\rho k} \right\rfloor^{1/\rho k} \sum_i \beta k_i = 1 \qquad k = 1, 2 \tag{8}$$

[8] Two other flexible forms based on Laurent series expansions are the Minflex Laurent Generalized Leontief and the Minflex Laurent translog (see Barnett, 1983; Barnett and Lee, 1985; Barnett *et al*. (1985, 1987)).

The share equations are

$$w_{it} = EP1_i(1 - Z_t) + EP2_i Z_t + u_{it} \qquad \text{where } EPk_i = \frac{\partial \ln pk}{\partial \ln p_i} \tag{9}$$

where Z is calculated from Equation 7 of Cooper and McLaren (1996). Conditions for effective global regularity for the systematic part of the share equations over the region $c > kP1$ require $\beta\kappa_i \geq 0$, $\rho k \leq 1$, $0 \leq \eta \geq 1$, $\mu \geq -1$ for $i = 1, \ldots, n$, and $k = 1, 2$.

3.3. Globally Flexible Functional Forms

As already pointed out, the functional forms considered so far are capable of approximating an arbitrary function locally at a single point in a delta neighbourhood of an often small but unknown size. A more general approach to approximating the true data-generating function is to use functional forms that have global properties. The idea behind these semi-non-parametric (SNP) functions is to expand the order of the series expansion, as the sample size increases, until the SNP function converges asymptotically to the true data-generating process and therefore to the true elasticities of substitution. Two such functional forms in general use are the Fourier flexible functional form and the Asymptotically Ideal Model (AIM). Monte Carlo studies of Fleissig *et al*. (1997), Terrell (1995) and Chalfant and Gallant (1985) show that the regularity region of the AIM and Fourier are much larger than that of the GL and BTL.

The Fourier Model

A way to obtain global flexibility—and gain some generality at the same time—is to estimate a demand system based on the classical Fourier sine/cosine series expansion of the reciprocal of the indirect utility function. Following Gallant (1981), the Fourier flexible form approximation of the indirect utility function may be written as

$$h_k(v) = u_0 + b'v + \frac{1}{2}v'Cv + \sum_{\alpha=1}^{A}\left(u_{0\alpha} + 2\sum_{j=1}^{J}[u_{j\alpha}\cos(jk'_\alpha v) - w_{j\alpha}\sin(jk'_\alpha v)]\right) \tag{10}$$

in which $C = -\sum_{\alpha=1}^{A} u_{0\alpha}k_\alpha k'_\alpha$ and v is a vector of the expenditure normalized prices. Here k_α is called a 'multi-index' and b, u, and w are the parameters to be estimated. The parameters are homogeneous of degree zero and the normalization $\sum b_i = 1$ is imposed. The empirical problem is to choose A and J to determine the length and degree of the approximation.

The Asymptotically Ideal Model (AIM)

Using Gallant's (1981) framework, Barnett and Jonas (1983) and Barnett and Yue (1988) developed the globally flexible AIM from the Müntz–Szatz series expansion of

$$h_k(v) = \alpha_0 + \sum_{k=1}^{K}\sum_{i=1}^{n} a_{ik}v_i^{\lambda(k)} + \sum_{k=1}^{K}\sum_{m=1}^{K}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ijkl}v_i^{\lambda(k)}v_j^{\lambda(m)}\right) + \cdots \tag{11}$$

Here $i \neq j$, $j \neq h$, $i \neq h$, $\lambda(k) = 2^{-k}$ for $k = 1, \ldots, \infty$ are the exponent set and $a_{ik}, a_{ijkl}, \ldots$ are the parameters to be estimated. As in the Fourier model, the order of the Müntz–Szatz series expansion is determined empirically.[9]

[9] For $K = 1$, the AIM is identically equal to the GL. Barnett and Yue (1988) provide an example of the share equations for $K > 1$.

4. THE US CONSUMPTION DATA

We use the quarterly private consumption expenditure data set constructed by Fleissig *et al.* (2000). This data set covers the period from 1960:1 to 1991:4 and is constructed using all of the disaggregated series from the US private consumption expenditure accounts. Prices and expenditure on nondurables and services are from the *National Income and Product Accounts* (NIPA). Since proper prices for durables are not available in the official statistics, the user cost of durables is calculated using Diewert's (1974) formula. Also, quarterly stock series for durables are constructed by extending the Campbell and Mankiw (1990) approach. All series are converted into per-capita terms using total population (see Fleissig *et al.*, 2000, for more details).

Fleissig *et al*. tested the hypothesis that this data set has been generated by a utility-maximizing agent. In particular, they used Varian's NONPAR program and showed that the consumption data are usually inconsistent with the Generalized Axiom of Revealed Preference (GARP).[10] They found, however, two subsamples over which there are no violations of GARP. These are 1960:1–1980:4 and 1981:4–1991:4. The aggregate consumption model is therefore consistent with the data over each of the two subsamples but not over the entire sample used in this paper. This suggests that we might expect no violations of the regularity conditions over each sub-sample but might encounter some violations when the entire data set is used.[11]

As this data set involves many categories of consumption, the estimation of a highly disaggregated demand system encompassing this many variables is econometrically intractable. This is especially important for SNP functions, since these functions can become very parameter intensive. Following Fleissig *et al*. (2000), we reduce the number of variables by constructing three sub-aggregates using Divisia aggregation methods. The Divisia index is used because Diewert (1976, 1978) shows that it is a superlative index that gives a second-order approximation to any arbitrary unknown aggregator function. The three sub-aggregates are motor vehicles, other durables, and non-durables and services (combined).

The data set warrants further discussion. A recent finding in the econometrics literature is that estimation and hypothesis testing critically depend on the integration and cointegration properties of the variables. In the context of linear demand systems, for example, Ng (1995) and Attfield (1997) test the null hypothesis of homogeneity (with respect to prices and nominal income) and show that this cannot be rejected once the time series properties of the data are imposed in estimation. They both use the AIDS model whose share equations are linear in the variables; this implies that testing for linear cointegration (in the spirit of Engle and Granger, 1987), and constructing a linear form of the error-correction model, is appropriate. In addition, Lewbel (1996) finds some evidence of non-stationarity for NIPA data using a similar approach. In our case, however, all models except the AIDS have share equations that are non-linear. As Granger (1995) points out, non-linear modelling of non-stationary variables is a new, complicated, and largely undeveloped area. We generally ignore this issue in this paper, keeping in mind that this is an area for future research. Nonetheless, as a proxy for non-stationarity we report unit root tests on the residuals for the estimated equations.

---

[10] NONPAR does a linear programming exercise to find if the data satisfy GARP and reports any violations. A violation occurs when a consumption bundle is revealed 'not preferred' in one period but then is bought in another period even though the bundle that was previously revealed preferred is a feasible choice.

[11] We note that for Varian's test the data are likely to pass GARP when income grows over the sample, which occurs for most of the data set. Clearly, the effectively globally regular functions are designed to accommodate such income growth.

Some other limitations of the data set used concern the quality and definition of the NIPA data recently discussed by Wilcox (1992) and Slesnick (1998). There is also a difference between purchases and the use of durables. We approximate the use of durables by using the user cost of durables and by assuming service flows are proportional to the stock of durables (as calculated by Fleissig *et al*., 2000). In addition, aggregate data have relatively less income variation than household data.

Lastly, Lewbel (1996) provides a new rationalization for aggregation across goods. He finds that the empirical assumptions regarding price movements of a group of NIPA goods appear to hold for a generalized composite commodity aggregation but are inconsistent with some assumptions of separable utility. Taking the results of Lewbel (1996) and ours together (we show separability), even though the data sets differ, we would argue that in our case the aggregation bias across goods may be relatively small.

5. ECONOMETRIC RESULTS

In the previous section, we have introduced eight systems of budget-share equations. McElroy (1987) shows that in estimation additive errors are preferable to multiplicative errors, so that we can write each system of budget share equations as $s_t = f(v_t, \theta) + e_t$. Further, since the budget shares sum to unity the disturbance covariance matrix is singular. Barten (1969) shows that the Maximum Likelihood estimates can be obtained by dropping any equation. All estimation is performed in International TSP 4.2 using the non-linear LSQ procedure.

In our initial tests the computed equation-by-equation Durbin–Watson statistics were low, suggesting significant positive serial correlation. We therefore assume a first-order autoregressive process (AR(1)) such that $e_t = \rho e_{t-1} + \varepsilon_t$ where $\rho = [\rho_{ij}]$ is a matrix of unknown parameters and ε_t is a non-autocorrelated vector disturbance term with constant covariance matrix.[12] The autocorrelation could be due to the effects of omitted dynamics (see e.g. Pollak and Wales, 1992), non-stationarity of prices (Lewbel, 1996), or the result of income effects that arise from aggregation across individual consumers (Stoker, 1986). As a proxy for possible omitted dynamics, demographic shifts, and deterministic non-stationarity, all models were estimated both without and with a time trend.

In our empirical work we have chosen to present two sub-periods of the data, 1960:1–1991:4 (the entire period) and 1960:1–1980:4. The latter period was chosen because, as discussed above, the underlying data were consistent with the General Axiom of Revealed Preference (GARP) for this sub-period. The second GARP-consistent period, 1981:4–1991:4, was judged to be too short for effective estimation to be possible. Note that all estimation includes an AR(1) correction, as already discussed, with symmetry, homogeneity and adding-up imposed.[13]

Before trying to compare these models, it is useful to look at unit root tests on the residuals from all of the models estimated. We use two alternative unit root testing procedures, to deal with anomalies that arise when the data are not very informative about whether or not there is a

[12] Following Berndt and Savin (1975), we assume that there is no autocorrelation across equations (i.e. assume that ρ is diagonal). As they point out, the autocorrelation coefficients for each equation must be identical for results to be invariant to which equation is deleted in the estimation.

[13] The actual coefficients produced by the various models, while interesting to specialists, are of no particular value in our comparisons of the models. Tables of these results are available from any of the authors. We do note, though, that each of the functional forms fits reasonably well.

Table I. Unit root tests 1960:1–1980:4[a]

	No time trend						Time trend					
	Eqn (1)		Eqn (2)		Eqn (3)		Eqn (1)		Eqn (2)		Eqn (3)	
	ADF	$Z(t_\alpha)$	ADF	$Z(t_\alpha)$	ADF	$Z(t_\alpha)$	ADF	$Z(t_\alpha)$	ADF	$Z(t_\alpha)$	ADF	$Z(t_\alpha)$
GL	0.000	0.000	0.000	0.000	0.230	0.000	0.000	0.000	0.000	0.000	0.230	0.000
BTL	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.000
AIDS	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.999	0.689
LAUR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
QUAID	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.999	0.877
GEF	0.383	0.000	0.373	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
AIM	0.044	0.000	0.072	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000
FFF	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.014	0.000

[a] p-values for the null hypothesis of a unit root ADF (Augmented Dickey–Fuller) and $Z(t_\alpha)$ (Phillips–Perron test).

unit root. In particular Table I reports p-values [based on the response surface estimates given by MacKinnon (1994)] for the augmented Dickey–Fuller (ADF) test (see Dickey and Fuller, 1981), and the non-parametric, $Z(t_\alpha)$ test of Phillips (1987) and Phillips and Perron (1988).

For the ADF test, the optimal lag length is taken to be the order selected by the Akaike Information Criterion (AIC) plus 2—see Pantula *et al.* (1994) for details regarding the advantages of this rule for choosing the number of augmenting lags. The $Z(t_\alpha)$ test is done with the same Dickey–Fuller regression variables, using no augmenting lags. Based on the p-values for the ADF and $Z(t_\alpha)$ unit root tests, the null hypothesis of a unit root is generally rejected for both sets of residuals.

Interpreting the results from estimating a system of non-linear equations is complicated when residuals appear to be stationary and when an AR(1) correction term is included. When estimating a linear system of equations, if the residuals are stationary, then the demand equations may be cointegrated. To obtain precise parameter estimates of the cointegrating vector when the variables are linear, Attfield (1997) and Ng (1995) suggest using the DOLS method of Stock and Watson (1995) or the FMOLS approach of Phillips (1995). Since these methods cannot be applied to a non-linear system of equations, we cannot modify the estimation procedure to adjust for the potential estimation bias in our tests. In addition, for a linear system of equations, if the AR(1) correction provides an approximation to first differencing the data, then the equations are misspecified unless an error-correction term is also included. Since the unit root tests have low power in distinguishing between a unit root and a near unit root, and our systems of equations are non-linear, we cannot determine if they are misspecified or not. While any of our inferences based on the residuals may be reliable since the residuals appear stationary, estimates of the standard errors for the parameters and forecasts may be imprecise. However, since the asymptotically flexible forms give a global approximation to the data function at all data points, the potential bias of the standard errors for these functions may be relatively small.

There is a tradeoff between estimating a system of linear or non-linear equations. If the data-generating function is non-linear, and this appears to be the case for our data set, then estimates from a linear approximation may be biased. If the variables are cointegrated then results from our non-linear models may be imprecise. Future research estimating non-linear models with data that may be non-stationary may consider bootstrapping or a Monte Carlo simulation, but such an investigation goes well beyond the scope of the present paper.

Table II. Performance statistics

LOCALLY FLEXIBLE FORMS

	No time trend						Time trend					
	1960:1–1991:4			1960:1–1980:4			1960:1–1991:4			1960:1–1980:4		
	Adj. R^2	LM[a]	MSE	Adj. R^2	LM[a]	MSE	Adj. R^2	LM[a]	MSE	Adj. R^2	LM[a]	MSE
GL Eqn (1)	0.952	24.301	0.000065	0.953	17.516	0.000080	0.996	13.319	0.000005	0.998	9.158	0.000003
GL Eqn (2)	0.970	17.424	0.000188	0.967	13.899	0.000250	0.996	10.717	0.000004	0.998	9.545	0.000003
GL Eqn (3)	0.931	24.284	0.000188	0.935	16.358	0.000250	0.993	6.749	0.000001	0.994	4.803	0.000001
BTL Eqn (1)	0.998	21.676	0.000004	0.998	9.630	0.000003	0.998	30.198	0.000003	0.998	29.661	0.000003
BTL Eqn (2)	0.998	3.354	0.000002	0.998	2.132	0.000002	0.999	15.169	0.000002	0.999	12.201	0.000002
BTL Eqn (3)	0.983	29.272	0.000003	0.978	11.034	0.000004	0.985	29.351	0.000003	0.982	8.391	0.000004
AIDS Eqn (1)	0.998	4.420	0.000003	0.998	7.503	0.000003	0.997	6.737	0.000003	0.998	4.531	0.000003
AIDS Eqn (2)	0.993	10.556	0.000010	0.994	11.126	0.000010	0.994	8.292	0.000008	0.994	15.181	0.000006
AIDS Eqn (3)	0.958	29.212	0.000014	0.952	35.484	0.000012	0.957	19.497	0.000015	0.950	25.443	0.000035

EFFECTIVELY GLOBALLY REGULAR

	No time trend						Time trend					
	1960:1–1991:4			1960:1–1980:4			1960:1–1991:4			1960:1–1980:4		
	Adj. R^2	LM[a]	MSE	Adj. R^2	LM[a]	MSE	Adj. R^2	LM[a]	MSE	Adj. R^2	LM[a]	MSE
Laur Eqn (1)	0.999	13.394	0.000002	0.999	13.612	0.000001	0.999	13.944	0.000003	0.999	8.110	0.000002
Laur Eqn (2)	0.999	13.656	0.000001	0.999	20.844	0.000001	0.999	9.808	0.000001	0.999	5.748	0.9×10^{-6}
Laur Eqn (3)	0.990	40.731	0.000001	0.998	20.195	0.000001	0.998	35.771	0.3×10^{-6}	0.998	15.967	0.3×10^{-6}
QAIDS Eqn (1)	0.997	9.541	0.000003	0.997	11.376	0.000003	0.998	9.323	0.000003	0.980	7.144	0.000003
QAIDS Eqn (2)	0.993	23.520	0.000002	0.991	25.186	0.000002	0.997	18.175	0.000002	0.979	9.325	0.000002
QAIDS Eqn (3)	0.954	6.873	0.000001	0.955	6.727	0.000001	0.976	65.753	0.000023	0.964	25.404	0.000170
GEF Eqn (1)	0.989	10.477	0.000002	0.995	9.654	0.000001	0.994	11.357	0.000001	0.997	7.574	0.000001
GEF Eqn (2)	0.997	9.514	0.000042	0.998	7.564	0.000004	0.998	8.235	0.000022	0.999	5.664	0.000003
GEF Eqn (3)	0.987	16.591	0.000034	0.996	14.351	0.000003	0.998	12.241	0.000018	0.998	9.276	0.000001

ASYMPTOTICALLY GLOBALLY FLEXIBLE FORMS

	No time trend						Time trend					
	1960:1–1991:4			1960:1–1980:4			1960:1–1991:4			1960:1–1980:4		
	Adj. R^2	LM[a]	MSE	Adj. R^2	LM[a]	MSE	Adj. R^2	LM[a]	MSE	Adj. R^2	LM[a]	MSE
AIM Eqn (1)	0.999	6.255	0.0000004	0.999	5.788	0.8×10^{-7}	0.999	9.996	0.000001	0.999	3.079	0.000001
AIM Eqn (2)	0.999	16.392	0.0000002	0.999	16.601	0.5×10^{-9}	0.999	8.024	0.000005	0.999	7.744	0.6×10^{-7}
AIM Eqn (3)	0.995	13.131	0.0000542	0.999	5.421	0.7×10^{-9}	0.998	33.949	0.3×10^{-6}	0.999	8.197	0.2×10^{-7}
FFF Eqn (1)	0.999	13.547	0.0002653	0.999	8.070	0.0001795	0.998	9.176	0.000340	0.998	7.840	0.000025
FFF Eqn (2)	0.999	20.269	0.0001210	0.999	10.910	0.0000859	0.999	8.787	0.000138	0.999	6.645	0.000009
FFF Eqn (3)	0.997	22.711	0.0000007	0.998	9.113	0.0000003	0.996	23.509	0.8×10^{-6}	0.998	10.239	0.4×10^{-6}

[a] Lagrange multiplier test for autocorrelation.
Notes:
Equation (1) is for motor vehicles.
Equation (2) is the aggregate of other durables.
Equation (3) is the aggregate of nondurables and services.

To begin the process of comparison, Table II reports performance statistics for all the functional forms. We test for autocorrelation using the Lagrange multiplier test because lagged endogenous variables are used. To perform the Lagrange multiplier test, the residuals from each share equation are regressed on the exogenous variables (expenditure normalized prices) and the lagged residuals from the *ith* share equation $\hat{e}_{it} = \sum_{j=1}^{4} \gamma_j v_j + \sum_{j=1}^{p} \rho_i \hat{e}_{it-j}$. The Lagrange multiplier test is distributed as Chi-square with p degrees of freedom. Failing to reject the null, $H_0 : \rho_1 = \rho_2 = \ldots = \rho_p = 0$, indicates no autocorrelation. Four lagged values are included since quarterly data are used; this gives a Chi-square test-statistic of 9.49 at the 5% level. The Lagrange multiplier statistics (shown as LM* in the table) indicate little or no serial correlation for the GEF, QUAIDS and both globally flexible forms. There is still some evidence of serial correlation for the other functional forms. The adjusted R^2 calculations are generally high for all functional forms and thus provide little information in distinguishing between specifications; similarly, the mean square error (MSE) is always very small in all of these tests.

No matter what degree of flexibility one adopts, there exists the possibility that regularity conditions will be rejected, implying rejection of the theory and/or the data, other things being equal. This is especially likely with local approximations since the frequency of such violations is likely to rise away from the point of approximation, but less likely for either the effectively globally flexible forms or the globally flexible forms. Even so, the more flexible functional forms are not immune to this problem for, after all, the data are bearing a heavy load of assumption no matter what technique is employed. In addition to data problems, violations of regularity can be attributed to factors such as omitted demographics, aggregation bias, and mis-specified dynamics.

Following Christensen *et al.* (1975), we have restricted the parameters of the estimated equations, as part of the maintained hypothesis, to satisfy adding-up, homogeneity, and symmetry. But we can test for the satisfaction of the theoretical restrictions that are not part of the maintained hypothesis at each point. These restrictions are non-negativity, monotonicity, and the curvature conditions on the indirect utility functions.[14] There are no violations of the non-negativity and monotonicity restrictions for all functional forms so the only regularity tests reported in Table III are the results of the quasi-convexity tests.

Table III shows the percentage of violations over the data space for each of the models. As is readily apparent, all functional forms have the same or fewer violations over the GARP consistent period than over the total period. This is the expected result. Furthermore, the BTL, AIDS and Laurent models have relatively more violations of concavity than the GL, GEF, QUAIDS, AIM, and Fourier models.[15] For all samples and whether a time trend is included or not, the GEF and QUAIDS models always had no violations of concavity. Recall that the regular region for these models grows as income grows, which probably occurred over much of the sample even though GARP is rejected over the entire sample. Thus, for this data set, the GEF and QUAIDS models find the data consistent with rational consumer behaviour even though the data failed the GARP test. On net, on the basis of this table, it appears that the QUAIDS, GEF, AIM and Fourier are the better models.

In our tests, we use standard information criteria for comparing the models. Table III (Part B) reports the Akaike Information Criterion (AIC) and Bayesian-Schwarz Information Criterion (SIC) for these tests. These information criteria, for which a low value is desirable, show that the GEF is the clear winner but that the differences among the models often are small. This is an important

[14] Non-negativity requires that the values of the fitted demand functions be non-negative ($x_i \geq 0, \forall_i$). Monotonicity requires that the indirect utility function be monotonically decreasing and this can be checked by direct computation of the values of the gradient vector of the estimated indirect utility function. Finally, the curvature conditions require quasi-convexity of the indirect utility function; these can be checked, provided the monotonicity conditions hold, by direct computation from the utility function.

[15] Fisher (1992) also found some violations of quasi-convexity using monetary data and the Fourier Flexible Form.

Table III. Eight flexible functional forms compared

(A) Quasi-convexity tests—percentage of violations

	No time trend		Time trend	
	1960:1–91:4	1960:1–80:4	1960:1–1991:4	1960:1–1980:4
GL	0.08	0.06	0.09	0.06
BTL	0.16	0.13	0.18	0.15
AIDS	0.18	0.14	0.20	0.22
Full Laurent	0.27	0.16	0.25	0.18
QUAIDS	0.00	0.00	0.00	0.00
GEF	0.00	0.00	0.00	0.00
AIM	0.04	0.00	0.05	0.00
Fourier	0.12	0.05	0.08	0.04

(B) Information criteria

	No time trend				Time trend			
	1960:1–1991:4		1960:1–1980:4		1960:1–1991:4		1960:1–1980:4	
	AIC	SIC	AIC	SIC	AIC	SIC	AIC	SIC
GL	1042.8	1042.9	668.1	668.3	1290.8	1290.9	862.3	862.5
BTL	1294.1	1294.3	839.5	839.7	1314.6	1314.8	860.6	861.0
AIDS	1209.3	1209.5	776.5	776.7	1213.3	1213.5	778.2	778.4
LAUR	1478.5	1478.6	978.6	978.8	1510.0	1510.1	994.1	994.3
QUA	1205.4	1205.6	836.9	837.0	1314.3	1314.4	803.6	803.7
GEF	994.6	993.8	789.4	788.5	948.2	948.0	778.2	777.6
AIM	1481.8	1482.4	967.6	968.4	1545.5	1546.1	989.9	990.6
FFF	1429.5	1429.9	957.7	958.2	1386.5	1386.9	900.8	901.3

result because then one might be more likely to prefer the effectively global forms over the AIM and FFF since the latter are more parameter intensive than the other functions.[16]

Next we provide a simple way of comparing how the eight functional forms measure income responses by examining how the models fit the three smallest and largest values over the estimated sample (1960–80); this appears in Table IV. We test to see if the fitted value is statistically significantly different from the realized value for the Wald statistic using the delta method. Adopting the 5% level of significance, we are looking for large p-values as evidence in favour of a particular specification. Accordingly, for the 36 'observations' in the table (taking 'time' and 'non-time' as separate tests), we have 23/36 instances in which the QUAIDS model successfully fit the extreme observations. This was the best performance, but the two other effectively global models (the Laurent and the GEF) did almost as well, as did the AIM model. We are not surprised at the performance of the QUAIDS model, since this rank 3 model was designed to deal with non-linear Engel curves.

[16] Theoretical results imply that both the AIM and Fourier will approximate the true data-generating function asymptotically, but the results can be different in small samples. On the other hand, the effectively globally regular GEF and QUAIDS specifications are not asymptotically globally regular so that both asymptotic and small sample results may differ.

Table IV. *P*-values for three smallest and largest expenditures

			Smallest								
			1961:1			1961:2			1962:1		
			sh1	sh2	sh3	sh1	sh2	sh3	sh1	sh2	sh3
GL	No	time	0.104	0.000	0.000	0.000	0.000	0.000	0.343	0.320	0.299
GL		time	0.990	0.003	0.000	0.926	0.027	0.015	0.166	0.000	0.000
BTL	No	time	0.216	0.000	0.000	0.000	0.000	0.000	0.084	0.000	0.000
		time	0.000	0.000	0.000	0.898	0.000	0.000	0.044	0.504	0.196
AIDS	No	time	0.418	0.000	0.000	0.573	0.000	0.000	0.059	0.014	0.000
AIDS		time	0.752	0.000	0.022	0.013	0.000	0.000	0.968	0.473	0.000
LAUR	No	time	0.988	0.223	0.084	0.329	0.031	0.428	0.692	0.003	0.000
LAUR		time	0.521	0.830	0.065	0.549	0.409	0.893	0.801	0.211	0.003
QAIDS	No	time	0.885	0.432	0.112	0.111	0.432	0.253	0.832	0.005	0.003
QAIDS		time	0.886	0.568	0.213	0.349	0.123	0.451	0.933	0.007	0.015
GEF	No	time	0.362	0.254	0.524	0.135	0.432	0.563	0.002	0.000	0.023
GEF		time	0.521	0.830	0.065	0.549	0.409	0.893	0.801	0.211	0.003
FFF	No	time	0.934	0.414	0.419	0.082	0.361	0.056	0.188	0.000	0.005
FFF		time	0.011	0.000	0.223	0.000	0.000	0.014	0.006	0.000	0.027
AIM	No	time	0.061	0.031	0.657	0.101	0.084	0.478	0.005	0.000	0.780
AIM		time	0.208	0.134	0.682	0.134	0.120	0.345	0.001	0.000	0.203
			Largest								
			1980:4			1980:1			1980:2		
			sh1	sh2	sh3	sh1	sh2	sh3	sh1	sh2	sh3
GL	No	time	0.000	0.000	0.000	0.043	0.000	0.001	0.000	0.000	0.000
GL		time	0.000	0.000	0.000	0.019	0.026	0.000	0.000	0.000	0.809
BTL	No	time	0.000	0.000	0.000	0.001	0.000	0.902	0.325	0.989	0.546
		time	0.000	0.000	0.008	0.317	0.000	0.035	0.549	0.612	0.380
AIDS	No	time	0.000	0.001	0.000	0.000	0.556	0.000	0.938	0.802	0.876
AIDS		time	0.000	0.006	0.000	0.000	0.396	0.000	0.801	0.550	0.000
LAUR	No	time	0.000	0.000	0.358	0.000	0.245	0.000	0.965	0.002	0.000
LAUR		time	0.000	0.000	0.000	0.002	0.002	0.021	0.849	0.850	0.894
QAIDS	No	time	0.000	0.005	0.032	0.000	0.643	0.003	0.966	0.533	0.333
QAIDS		time	0.000	0.012	0.000	0.059	0.444	0.041	0.941	0.125	0.414
GEF	No	time	0.000	0.000	0.124	0.000	0.108	0.003	0.086	0.005	0.000
GEF		time	0.000	0.000	0.000	0.009	0.071	0.022	0.051	0.631	0.000
FFF	No	time	0.000	0.000	0.016	0.000	0.001	0.000	0.062	0.000	0.000
FFF		time	0.000	0.000	0.000	0.000	0.068	0.000	0.000	0.934	0.000
AIM	No	time	0.000	0.000	0.035	0.001	0.262	0.000	0.814	0.216	0.006
AIM		time	0.000	0.000	0.000	0.005	0.022	0.003	0.114	0.462	0.009

Notes: The null is rejected at the 5% level for values less than 0.05. sh1 = share of motor vehicles, sh2 = other durables aggregate, sh3 = aggregate of non-durables and services.

The final issue, for our data set, is how to chose among the GEF, QUAIDS and the two global models, the AIM and the Fourier. A method of comparing the informational content of the parameter estimates is to look at the behaviour of the output of the models in the form of the elasticities of substitution; we will do this in Section 6. Here we are looking for inconsistencies in the pattern of elasticities from one model to the next. Our final way of comparison, discussed in Section 7, is through out-of-sample forecasts.

6. MORISHIMA AND EXPENDITURE ELASTICITIES OF SUBSTITUTION

Although the Allen elasticity of substitution (AES) has been used widely to study substitution behavior and structural instability, Blackorby and Russell (1981, 1989) have shown that the AES is quantitatively and qualitatively ambiguous. In fact, when there are more than two variables, the Morishima elasticity of substitution (MES) is a generally better measure of the substitution elasticity.[17] The Morishima elasticity of substitution, $\sigma^{m}{}_{ij}$, is defined (Blackorby and Russell 1981, 1989) as follows:

$$\sigma_{ij}^{m} = s_i(\sigma_{ij}^{a} - \sigma_{li}^{a})$$

It categorizes goods as complements ($\sigma^{m}{}_{ij} < 0$) if an increase in the price of j causes x_i/x_j to decrease. If $\sigma^{m}{}_{ij} > 0$, goods are Morishima substitutes.

We are interested, in this paper, in the relations among three types of consumption goods. To this point it has not been possible to dig into the economic questions, but with a measure of the elasticity of substitution, this is now possible. The types of goods we are working with are (1) motor vehicles, (2) other durables, and (3) non-durables and services. We would expect each of these broad categories to be a substitute for each of the others and we might expect the substitution to be stronger between 1 and 2 than between, say, 1 and 3. In what follows, then, we will try to balance our interest in model comparison with the underlying economics.

The Morishima elasticities between motor vehicles (1) and other durables (2) for all eight models excluding a time trend are displayed in Figure 1. Note that this is the Morishima elasticity calculated by varying the price of motor vehicles.[18]

Here the asymptotically globally flexible AIM and Fourier models show elasticities in the range between 0.25 and 1, while the other four (local) approximations show lower substitution, with the

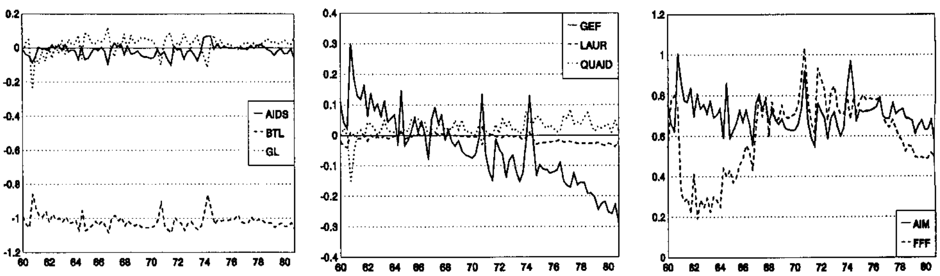

Figure 1. Morishima elasticities for motor vehicles and other durables (price of motor vehicles varying)

[17] The AES between goods i and j, $\sigma^{a}{}_{ij}$, is traditionally computed from the cost function as

$$\sigma_{ij}^{a} = \frac{C(u,P)C_{ij}(u,P)}{c_i(u,P)c_j(u,P)}$$

It categorizes goods as complements if an increase in the price of good j causes a decreased consumption of good $i(\sigma^{a}{}_{ij} < 0)$. If $\sigma^{a}{}_{ij} > 0$, goods are Allen substitutes.

[18] In Table V below, we discuss the precision of the estimates by displaying the percentage of statistically significant elasticities.

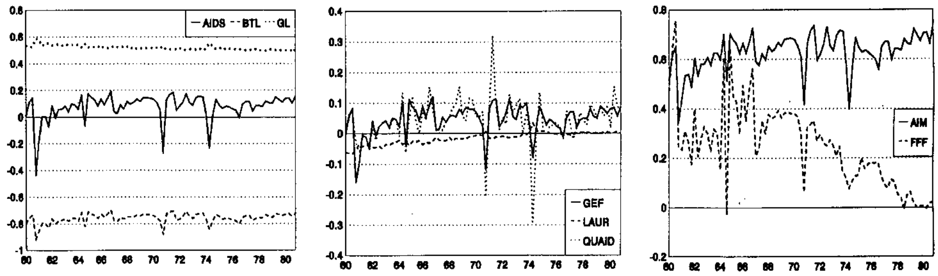

Figure 2. Morishima elasticities for other durables and non-durables (price of non-durables varying)

basic translog model showing complementarity (around-1.0). All the remaining flexible models have a number of cases in which complementarity is exhibited. We believe these results show how model-specific results can be in this literature. Of course, consumer theory does not rule out complementarity, so we can hardly claim that the results for the Fourier or AIM model are correct. Nevertheless, most researchers expect substitution rather than complementarity for such broad aggregates. It is also noteworthy in Figure 1 that the two asymptotically globally flexible models show considerably more variation in the elasticities of substitution. Again this is a result that appears to be model-specific. While we have no *a priori* view as to what is appropriate here, it is worth emphasizing that obtaining greater flexibility with models that are capable of estimating the elasticity to an arbitrary degree of accuracy *globally*, suggests that the underlying utility surfaces are in fact non-linear and perhaps highly so. This might suggest that studies of aggregate consumption would do well to maintain this level of disaggregation, since weak and/or variable substitution elasticities among the components make for poorly behaved aggregates, on the whole.

In Figure 2 we pursue these questions with respect to the relations between other durables (2) and non-durables and services (3). Our main interest is in investigating what are likely to be the strongest links in the chain, and we anticipate somewhat lower substitution in this case. This appears to be true for the AIM and Fourier models, both of which show lower substitution than in Figure 1. For the other models, the results are ambiguous, on the whole. Figure 2, further, again shows that the globally functional forms (the AIM and Fourier models) show positive elasticities (substitution) and typically greater variability of the Morishima parameter (we are varying the price of ‘other durables’ in this experiment). It is noticeable, in fact, that all of the series show a ‘cyclical’ pattern, much of which seems related to cyclical activity in the US economy. We hypothesize that business cycle turning points produce relatively large changes in user costs that, in turn, interact with the highly non-linear utility function of the aggregate consumer. The asymptotically globally flexible functional forms seem more adept at picking up this behaviour than the local functions studied here. While this is merely an interpretation, it is worth emphasizing that the QUAIDS, GEF, and AIM specifications exhibited *no* failures of the regularity conditions at any points in the data (and the Fourier failed 6% of the time), and that this data set passed the tests for conformity with the General Axiom of Revealed Preference and weak separability. Notice, finally, that the basic translog model continues to exhibit complementarity between these two commodity bundles.

Pursuing the topic of the expenditure elasticities, we first look at those for non-durables in Figure 3. We concentrate on expenditure elasticities from the GEF, QUAIDS, AIM and Fourier flexible forms since these appear to be the functions that give the best approximations on these

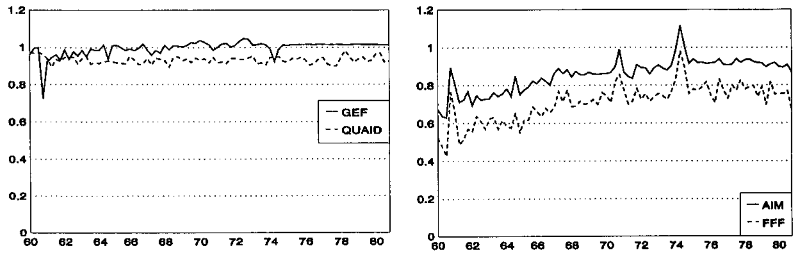

Figure 3. Expenditure elasticities for non-durables

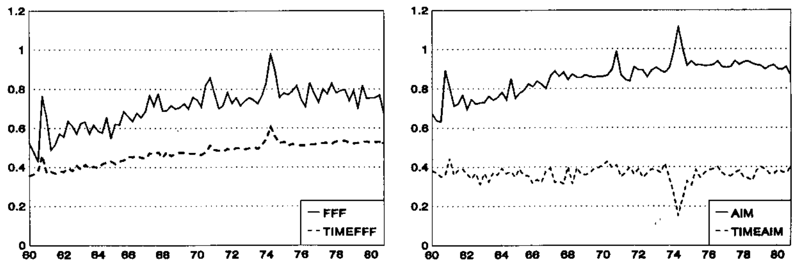

Figure 4. Expenditure elasticities and the effect of time trend

data. These expenditure elasticities, calculated without a time trend, range from 0.4 to 1.1 with the AIM and Fourier estimates, for the most part, slightly smaller estimates. We have no firm priors about these elasticities, although we find the fact that the estimated expenditure elasticities are often less than unity for nondurables a reasonable finding. We note that the two globally flexible functions show rising expenditure elasticities through the 1960s, while the two effectively global functions tested do not. We suspect the rising elasticities are possibly correct, as per capita incomes rose considerably during this period.

Finally, Figure 4 is used to analyse the affect of including a time trend and shows the expenditure elasticities for the AIM and Fourier functions. It is apparent that including a time trend lowers estimates of expenditure elasticities and may even produce considerably less variability in the substitution elasticities. It also removes the upward drift in the AIM model, but not in the Fourier. It is worth noting that in all these graphs there are sharp spikes at the start of the recession of 1973–5.

7. FORECAST RESULTS

A final way to evaluate contrasting models of the same phenomenon is to compare their out-of-sample forecasting performance. It appears as though the two asymptotically globally flexible models provide the best approximations for the data used in this paper, but it is common among such studies to find that simpler methods sometimes forecast as well as more complicated methods. Of course, our main interest is in comparing the effectively global and the asymptotically global specifications but, as will be readily apparent, we cannot make a clear case. There are actually numerous ways to compare the forecasting capabilities of econometric models. We follow Mathews and Diamantopoulos (1994) who suggest four measures based on an extensive evaluation of many methods for evaluating forecasts. They propose using the following metrics: the average absolute percentage error, root mean square error, mean absolute error, and *R*-Square. All eight functional forms were used to predict the shares for 1, 2, and 3 years ahead starting at 1981:1, using the AR(1) correction. That is, the eight functional forms were estimated over the period 1960:1–1980:4 and then forecasted to 1983:4 for one-, two-, and three-period intervals. The results are displayed in Table V, aggregated for each year in the forecasts.

Table V. Evaluating Forecasts

			Mean Absolute Error			Root Mean Square Error		
(A) No time	trend	Year	share1	share2	share3	share1	share2	share3
	GL	1	0.000032	0.000024	0.000015	0.000206	0.000156	0.000090
		2	0.000079	0.000049	0.000037	0.000348	0.000213	0.000182
		3	0.000115	0.000068	0.000054	0.000423	0.000246	0.000217
	AIM	1	0.000063	0.000044	0.000031	0.000309	0.000230	0.000154
		2	0.000133	0.000088	0.000059	0.000464	0.000309	0.000234
		3	0.000206	0.000133	0.000087	0.000585	0.000376	0.000277
	FFF	1	0.000045	0.000058	0.000031	0.000230	0.000290	0.000167
		2	0.000106	0.000100	0.000076	0.000381	0.000365	0.000299
		3	0.000150	0.000139	0.000099	0.000450	0.000411	0.000323
Time	trend	Year	share1	share2	share3	share1	share2	share3
	LAUR	1	0.000041	0.000025	0.000019	0.000212	0.000162	0.000097
		2	0.000085	0.000049	0.000079	0.000295	0.000210	0.000309
		3	0.000148	0.000068	0.000129	0.000412	0.000236	0.000394
	AIM	1	0.000083	0.000055	0.000027	0.000402	0.000276	0.000128
		2	0.000162	0.000124	0.000063	0.000559	0.000455	0.000212
		3	0.000207	0.000162	0.000092	0.000604	0.000490	0.000252
	FFF	1	0.000035	0.000040	0.000028	0.000168	0.000188	0.000175
		2	0.000078	0.000085	0.000040	0.000261	0.000280	0.000186
		3	0.000106	0.000104	0.000052	0.000298	0.000297	0.000196
(B) No time	trend		Average Absolute Percentage Error			Adjusted R-Squared		
		Year	share1	share2	share3	share1	share2	share3
	LAUR	1	0.003487	0.003002	0.001259	0.998597	0.999426	0.997227
		2	0.004235	0.002722	0.001908	0.998829	0.999723	0.989396
		3	0.004042	0.002416	0.001937	0.999217	0.999787	0.993006
	AIM2	1	0.004123	0.002585	0.001923	0.998233	0.999561	0.986208
		2	0.004419	0.002609	0.001814	0.998667	0.999771	0.990851
		3	0.004443	0.002688	0.001787	0.998953	0.999798	0.993725
	FFF	1	0.002928	0.003470	0.001912	0.998633	0.999235	0.998173
		2	0.003502	0.002965	0.002338	0.998775	0.999575	0.984512
		3	0.003225	0.002807	0.002043	0.999158	0.999624	0.989942
Time	trend	Year	share1	share2	share3	share1	share2	share3
	LAUR	1	0.002730	0.001528	0.001234	0.998720	0.998947	0.999306
		2	0.002834	0.001468	0.002450	0.999279	0.999467	0.994677
		3	0.003193	0.001382	0.002666	0.999446	0.999333	0.995759
	AIM	1	0.005478	0.003323	0.001691	0.997326	0.998673	0.991954
		2	0.005370	0.003693	0.001938	0.998615	0.998949	0.997053
		3	0.004454	0.003276	0.001895	0.998168	0.998319	0.997633
	FFF	1	0.002310	0.002390	0.001784	0.999379	0.999960	0.991722
		2	0.002592	0.002539	0.001252	0.999480	0.999962	0.995693
		3	0.002282	0.002112	0.001081	0.999555	0.999741	0.996174

While this is not a complete tabulation, it covers the best forecasting results for the eight models. Our selection criterion was to use the best two results in each category and then to include in the table those functions that dominated in this comparison. In Part A of the table, the GL, AIM, and Fourier produced an overwhelming percentage of the best results (11/12 for line 1, for example). When the time trend was included, the Laurent replaced the GL; the AIM and Fourier still came in well (and the three again produced 11/12 best results for line 1). In Part B, the Laurent, AIM and Fourier dominated, whether the time trend was included or not. Inescapably, the globally flexible AIM and Fourier dominate in these comparisons, with the effectively global Laurent not far behind.

To test if the forecasted value is statistically different from the realized value, we calculate p-values for the Wald statistic using the delta method. At the 5% level of significance, all models without a time trend fail to accept the null at over 50% of the quarters with the worst case of the GEF rejecting the null for all shares in all quarters. The results are similar for the models once a time trend is included but the Laurent model does particularly well (Table VI). Overall, no flexible form appears to be statistically better than the others in this list.

8. CONCLUSIONS

This paper evaluates and compares three types of flexible functional forms—locally flexible, effectively globally regular, and asymptotically globally flexible in terms of violations of regularity conditions, information criteria, performance in the tails of the expenditure space, substitution elasticities, and forecasts. All models are estimated using US aggregate consumption data that itself was found to be consistent with a well behaved utility function over much of the sample. As we shall explain in our summary, the global models, especially the QUAIDS, FFF, and AIM seem to have dominated on these tests.

The breakdown of the tests is as follows. All the models fit the data well, but a preference should be expressed for the more parametrically parsimonious functions; these are the GL and the AIDS models. Over the GARP consistent data set, quasi-concavity tests indicated that the QUAIDS, GEF, AIM, FFF, and GL did the best, with and without the time trend. The SIC and AIC tests favoured the GL, AIDS, GEF, and QUAIDS models, with the GEF doing the best. The Laurent, GEF, and particularly the QUAIDS model fit the extreme levels of expenditures best; the AIM model also did well in this test. Looking at substitution elasticities, we found those from the globally flexible AIM and FFF most plausible, with the GL least plausible. Finally, more support was obtained for the forecasting performance of the AIM, FFF, and Laurent models, although the GL was best in one category.

Across all tests, three specifications seem to stand out. These are the QUAIDS, FFF, and AIM models, the first being effectively global and the latter two being asymptotically global. The GEF model (also an effectively global specification) also did well on most tests. If nothing else, the importance of employing a globally regular model is convincingly demonstrated in this paper. We cannot say whether this should be achieved 'effectively' or 'asymptotically' on the evidence, but if one worries about parametric parsimony, then the effectively global methods (QUAIDS and GEF) might be preferable.

A particularly interesting result in this paper is that the LAURENT, QUAIDS, GEF and the globally flexible functions generally find that motor vehicles, other consumer durables, and non-durables generally are substitutes for each other. Finding substitution is a result most researchers

Table VI. *P*-values for one-to four-quarters-ahead forecasts[a]

No time trend

	GL			BTL			AIDS		
Year	sh1	sh2	sh3	sh1	sh2	sh3	sh1	sh2	sh3
81.1	0.000	0.000	0.000	0.000	0.045	0.529	0.709	0.000	0.000
81.2	0.000	0.921	0.016	0.000	0.000	0.195	0.000	0.046	0.107
81.3	0.000	0.000	0.000	0.000	0.001	0.067	0.000	0.000	0.777
81.4	0.007	0.194	0.378	0.000	0.000	0.000	0.000	0.080	0.026

	LAUR			QAIDS			GEF		
Year	sh1	sh2	sh3	sh1	sh2	sh3	sh1	sh2	sh3
81.1	0.170	0.736	0.017	0.000	0.180	0.001	0.000	0.007	0.000
81.2	0.000	0.000	0.058	0.000	0.000	0.071	0.000	0.000	0.001
81.3	0.149	0.000	0.002	0.040	0.000	0.000	0.003	0.001	0.008
81.4	0.001	0.000	0.569	0.000	0.000	0.000	0.000	0.000	0.000

	FFF			AIM		
Year	sh1	sh2	sh3	sh1	sh2	sh3
81.1	0.002	0.180	0.001	0.055	0.927	0.000
81.2	0.000	0.000	0.131	0.000	0.000	0.001
81.3	0.444	0.000	0.000	0.136	0.001	0.018
81.4	0.000	0.000	0.000	0.000	0.000	0.010

Time trend

	GL			BTL			AIDS		
Year	sh1	sh2	sh3	sh1	sh2	sh3	sh1	sh2	sh3
81.1	0.000	0.001	0.000	0.944	0.161	0.237	0.194	0.000	0.000
81.2	0.000	0.000	0.147	0.000	0.001	0.000	0.000	0.001	0.000
81.3	0.000	0.000	0.003	0.000	0.000	0.378	0.000	0.000	0.000
81.4	0.000	0.000	0.231	0.000	0.016	0.000	0.000	0.001	0.000

	LAUR			QAIDS			GEF		
Year	sh1	sh2	sh3	sh1	sh2	sh3	sh1	sh2	sh3
81.1	0.211	0.045	0.539	0.000	0.000	0.000	0.000	0.001	0.025
81.2	0.001	0.006	0.000	0.000	0.000	0.055	0.000	0.000	0.000
81.3	0.415	0.921	0.071	0.044	0.000	0.000	0.202	0.103	0.065
81.4	0.471	0.958	0.071	0.000	0.000	0.000	0.000	0.006	0.000

	FFF			AIM		
Year	sh1	sh2	sh3	sh1	sh2	sh3
81.1	0.001	0.002	0.000	0.000	0.001	0.013
81.2	0.000	0.000	0.436	0.000	0.000	0.000
81.3	0.017	0.000	0.000	0.137	0.202	0.174
81.4	0.000	0.000	0.000	0.000	0.006	0.000

[a]The null is rejected at the 5% level for values less than 0.05.
Note: sh1 = share of motor vehicles, sh2 = other durables aggregate, sh3 = aggregate of non-durables and services.

expect. In addition, in the better-fitting models we found sharp changes in these elasticities around recession periods (and occasionally at other times), results which if correct call into question the use of constant elasticity of substitution methods. The advantages of the global models are relatively greater when utility surfaces are highly nonlinear, when user costs fluctuate considerably over time, and (consequently) when estimated elasticities of substitution are significantly volatile (over time). Their relatively good performance compared to the local flexible forms in the period tested suggest that these observations have some merit. We conclude that the AIM model, closely followed by the Fourier, provides the most satisfactory results.

Finally, we note that the low income variation and the relatively large price variation often found in aggregate data may favor employing functional forms that allow for more price flexibility (such as the AIM and Fourier). Aggregate data may also bias results toward functions having simple Engel curves and/or low rank although our results show no evidence of this. In contrast, functions having more income flexibility relative to price flexibility may be preferred when using household-level data, such as the QUAIDS and GEF models. We think future research should examine how flexible forms perform using household-level data, particularly when there are many variables and a relatively small sample which may limit the use of the asymptotically globally flexible forms. Developing an asymptotically globally flexible QUAIDS or GEF model by adding Fourier series or a Muntz–Szatz expansion is another useful direction for future work.

ACKNOWLEDGEMENTS

We would like to thank Arie Kapteyn and the referees for comments that greatly improved this paper. Serletis acknowledges support from the Social Sciences and Humanities Research Council of Canada.

REFERENCES

Alston J, Foster KA, Green R. 1994. Estimating elasticities with the linear approximate ideal demand system: some Monte Carlo results. *Review of Economics and Statistics* **76**: 351–356.

Attfield CLF. 1997. Estimating a cointegrating demand system. *European Economic Review* **41**: 61–73.

Banks J, Blundell R, Lewbel A. 1997. Quadratic Engel curves and consumer demand. *Review of Economics and Statistics* **79**: 527–539.

Barnett WA. 1983. New indices of money supply and the flexible Laurent demand system. *Journal of Business and Economic Statistics* **1**: 7–23.

Barnett WA. 1985. The Minflex Laurent translog flexible functional form. *Journal of Econometrics* **30**: 33–44.

Barnett WA, Jonas A. 1983. The Muntz–Szatz demand system: an application of a globally well-behaved series expansion. *Economic Letters* **11**(4): 337–342.

Barnett WA, Lee YW. 1985. The global properties of the Minflex Laurent, generalized Leontief and translog flexible functional forms. *Econometrica* **53**: 1421–1437.

Barnett WA, Lee YW, Wolfe MD. 1985. The three-dimensional global properties of the Minflex Laurent, generalized Leontief, and translog flexible functional forms. *Journal of Econometrics* **30**: 3–31.

Barnett WA, Lee YW, Wolfe MD. 1987. The global properties of the two Minflex Laurent flexible functional forms. *Journal of Econometrics* **36**: 281–298.

Barnett WA, Yue P. 1998. Semiparametric estimation of the Asymptotically Ideal Model: the AIM demand system. In *Nonparametric and Robust Inference, Advances in Econometrics* 7, Rhodes G, Formby T (eds). JAI Press: Greenwich, CT; 229–252.

Barten A. 1969. Maximum likelihood estimates of a complete system of demand equations. *European Economic Review* **1**: 7–73.

Berndt ER, Savin NE. 1975. Estimation and hypothesis testing in singular equation systems with autoregressive disturbances. *Econometrica* **43**: 937–959.
Blackorby C, Russell RR. 1981. The Morishima elasticity of substitution, symmetry, constancy, separability, and its relationship to the Hicks and Allen elasticities. *Review of Economic Studies* **48**: 147–158.
Blackorby C, Russell RR. 1989. Will the real elasticity of substitution please stand up? A comparison of the Allen/Uzawa and Morishima elasticities. *American Economic Review* **79**: 882–888.
Buse A. 1994. Evaluating the linearized Almost Ideal Demand system. *American Journal of Agricultural Economics* **76**: 781–793.
Campbell J, Mankiw NG. 1990. Permanent income, current income, and consumption. *Journal of Business and Economic Statistics* **8**: 265–269.
Caves D, Christensen L. 1980. Global properties of flexible functional forms. *American Economic Review* **70**: 422–432.
Chalfant JA. 1987. A globally flexible Almost Ideal System. *Journal of Business Economic Statistics* **5**(2): 233–242.
Chalfant JA, Gallant AR. 1985. Estimating substitution elasticities with the Fourier costs function: some Monte Carlo results. *Journal of Econometrics* **28**: 205–222.
Christensen LR, Jorgenson DW, Lau LJ. 1975. Transcendental logarithmic utility functions. *American Economic Review* **65**: 367–383.
Cooper RJ, Mclaren KR. 1996. A system of demand equations satisfying effectively global regularity conditions. *Review of Economics and Statistics* **78**: 359–364.
Deaton A, Muellbauer J. 1980. An Almost Ideal Demand system. *American Economic Review* **70**: 312–326.
Dickey DA, Fuller WA. 1981. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica* **49**: 1057–1072.
Diewert WE. 1971. An application of the Shephard duality theorem: a generalized Leontief production function. *Journal of Political Economy* **79**: 481–507.
Diewert WE. 1974. Applications of Duality Theory. In *Frontiers of Quantitative Economics II*, Intriligator M, Kendrick D (eds). Amsterdam; North-Holland: 106–171.
Diewert WE. 1976. Exact and superlative index numbers. *Journal of Econometrics* **4**: 115–146.
Diewert WE. 1978. Superlative index numbers and consistency in aggregation. *Econometrica* **46**: 883–900.
Eastwood B. 1991. Asymptotic normality and consistency of semi-nonparametric regression estimators using an upwards F-test truncation rule. *Journal of Econometrics* **57**: 307–340.
Elbadawi IA, Gallant AR, Sousa G. 1983. An elasticity can be estimated consistently without a prior knowledge of functional form. *Econometrica* **51**: 1731–1752.
Engle RF, Granger CWJ. 1987. Cointegration and error correction: representation, estimation and testing. *Econometrica* **55**: 251–276.
Fisher D. 1992. Money-demand variability: a demand-systems approach. *Journal of Business and Economic Statistics* **10**: 143–152.
Fleissig AR, Hall AR, Seater JJ. 2000. GARP, separability, and the representative agent. *Macroeconomic Dynamics* Vol 4, No 3, September.
Fleissig AR, Kastens T, Terrell D. 1997. Semi-nonparametric estimates of substitution elasticities. *Economics Letters* **54**(3): July, 209–219.
Gallant AR. 1981. On the bias in flexible functional forms and an essentially unbiased form: the Fourier flexible form. *Journal of Econometrics* **15**: 211–245.
Gorman WM. 1959. Separable utility and aggregation. *Econometrica* **27**: 469–481.
Granger CWJ. 1995. Modelling nonlinear relationships between extended-memory variables. *Econometrica* **63**: 265–279.
Guilkey D, Lovell C. 1980. On the flexibility of the translog approximation. *International Economic Review* **21**: 137–147.
Guilkey D, Lovell C, Sickles C. 1983. A comparison of the performance of three flexible functional forms. *International Economic Review* **24**: 591–616.
Lewbel A. 1987a. Characterizing some Gorman–Engel curves. *Econometrica* **55**: 1451–1459.
Lewbel A. 1987b. Fractional demand systems. *Journal of Econometrics* **36**: 311–337.
Lewbel A. 1990. Full rank demand systems. *International Economic Review* **31**(2): May, 289–300.
Lewbel A. 1991. The rank of demand systems: theory and nonparametric estimation. *Econometrica* **59**: 711–730.

Lewbel A. 1995. Utility functions and global regularity of fractional demand systems. *International Economic Review* **36**(4): November, 943–961.

Lewbel A. 1996. Aggregation without separability: a generalized composite commodity theorem. *American Economic Review* **86**(3): June, 524–543.

Mathews BP, Diamantopoulos A. 1994. Towards a taxonomy of forecast error measures. *Journal of Forecasting* **13**: 409–416.

McElroy MB. 1987. Additive general error models for production, costs, and derived demand or share systems. *Journal of Political Economy* **95**: 737–757.

MacKinnon JG. 1994. Approximate asymptotic distribution functions for unit-root and cointegration tests. *Journal of Business and Economic Statistics* **12**: 167–176.

Ng S. 1995. Testing for homogeneity in demand systems when the regressors are nonstationary. *Journal of Applied Econometrics* **10**: 147–163.

Pantula SG, Gonzalez-Farias G, Fuller WA. 1994. A comparison of unit-root test criteria. *Journal of Business and Economic Statistics* **12**: 449–459.

Pashardes P. 1993. Bias in estimating the Almost Ideal Demand system with the Stone index approximation. *Economic Journal* **103**: 908–915.

Phillips PCB. 1987. Time series regression with a unit root. *Econometrica* **55**: 277–301.

Phillips PCB. 1995. Fully modified least squares and vector autoregression. *Econometrica* **62**: 1023–1078.

Phillips PCB, Perron P. 1988. Testing for a unit root in time series regression. *Biometrica* **75**: 335–346.

Pollak RA, Wales TJ. 1980. Comparison of the quadratic expenditure system and translog demand systems with alternative specifications of demographic effects. *Econometrica* **48**: 595–612.

Pollak RA, Wales TJ. 1992. *Demand System Specification and Estimation*. Oxford University Press: New York.

Ramajo J. 1994. Curvature restrictions on flexible functional forms: an application of the minflex Almost Ideal Demand system to the pattern of Spanish demand. *Journal of Business and Economic Statistics* **4**: 431–436.

Serletis A. 1988. Translog flexible functional forms and substitutability of monetary assets. *Journal of Business and Economic Statistics* **6**: 59–68.

Slesnik DT. 1998. Are our data relevant to the theory? The case of aggregate consumption. *Journal of Business and Economic Statistics* **16**: 52–61.

Stock JH, Watson MW. 1995. A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica* **61**(4): 783–820.

Stoker TM. 1986. Simple tests of distributional effects on macroeconomic equations. *Journal of Political Economy* **94**: 763–795.

Terrell D. 1995. Flexibility and regularity properties of the asymptotically ideal production model. *Econometric Reviews* **14**: 1–17.

Varian HR. 1982. The nonparametric approach to demand analysis. *Econometrica* **50**: 945–974.

Varian HR. 1983. Non-parametric tests of consumer behaviour. *Review of Economic Studies* **50**: 99–110.

Wilcox DW. 1992. The construction of U.S. consumption data: some facts and their implications for empirical work. *American Economic Review* **82**: 922–941.

