Model Specification Issues in Consumer Demand Systems Using United States Microdata[1]

by

Christopher J. Nicol
Department of Economics
University of Regina
Regina, Saskatchewan, S4S 0A2
CANADA
306-585-4182; FAX 306-585-4815
Internet: nicolc@elsie.econ.uregina.ca

October 3, 1995

[1]The support of the Social Sciences Research Council of Canada is gratefully acknowledged. I thank James Farley for outstanding research assistance. Any errors are my responsibility.

Abstract

A rank three demand system incorporating labour force participation, non-separability of demands from excluded goods and non-exact aggregation in income and household characteristics is estimated using United States Consumer Expenditure Survey microdata. Various models are estimated using ML and GMM methoods.

It is found that demands are not separable from labour force variables or age of head. The rank three requirement does not seem important here, however. Predictive power is extremely poor for GMM relative to ML estimated models, even though some variables appear to be non-exogenous. There is also evidence that errors in price variables can have important effects on estimationresults.

Keywords:

separability; rank three demands; labour supply; microdata.

1 Introduction

Recent research in applied demand analysis has focused on model specification issues. For example: whether demand systems are rank two or three (Fry and Pashardes, 1992 and Banks, Blundell and Lewbel, 1994); the separability of commodity demands from labour supply (Browning and Meghir, 1991 and Kaiser, 1993); and the role of household characteristics variables in demand models (Blundell, Pashardes and Weber, 1993 and Dickens, Fry and Pashardes, 1993). These studies indicate that each of the above features are important determinants of demand. However, one or more of these considerations have *not* been explicitly modelled. This is a matter for some concern. This was highlighted in Dagenais (1994), where the effects of multiple model specification errors were investigated. It was found that, when more than one model specification error was present, correcting for only one induced larger biases and inconsistencies than if one did not control for any of them.

In a previous paper (Nicol, 1995), Canadian cross-sectional microdata from eight Family Expenditure Surveys (FAMEX) were used to determine the relative importance of each of the foregoing modelling aspects. Some additional model features were also analysed, as was the practice of controlling for possible non-exogeneity of some explanatory variables. As a consequence, the models considered in that paper were estimated by maximum likelihood (ML) and the generalised method of moments (GMM). While there was evidence of non-exogeneity of some explanatory variables, controlling for this by GMM estimation did not necessarily improve matters. That is, a comparison of predictive power indicated that ML estimated prediction was typically superior to GMM estimated prediction.

Given the results in the above mentioned paper, it is of interest to explore whether these hold more generally, using other bodies of data. In this paper, the United States Consumer Expenditure Survey (CES) microdata are used to analyse these issues. As with the Nicol (1995) work, homogeneous groups of households are extracted from the CES data, covering the 1980–1992 survey years inclusive. Proportions of these data sets are used to conduct exploratory analysis of the type in Nicol (1995). This leaves most of the samples available for subsequent work, based on the findings with this baseline study.

Applied demand work with the CES data has not been conducted extensively, since there are no published price data from US Government sources reflecting inter-city or inter-regional differences in prices across the US at a point in time. This is in contrast to the Canadian case, where such data are published by Statistics Canada. Research which has used the CES has typically employed one cross-section of data, and assumed all agents in the sample face identical prices for different goods. These data have then been pooled with aggregate, time-series data, to identify the price parameters of the model. The work of Jorgenson, Lau and Stoker (1982) is an example of the use of the CES data in this way.

There is evidence in the US that all households across the country do not face identical prices at a point in time. However, without appropriate data, it is not possible to relax this assumption. In this paper, price data are constructed which reflect inter-city and inter-regional differences across the US at a point in time, *relative to the same US average and time of reference*. This permits the use of the cross-sectional data in a way where individual households located in different places are matched more closely to the prices they actually faced. This is the first time work of this kind has been done in an applied demand context for the US.

Using data published by the *American Chamber of Commerce Researcher' Association* (ACCRA), US Consumer Price Index (CPI) data are converted for a number of cities and states, yielding the appropriate kind of price data for a number of types of expenditures. These data are then used to estimate a variety of demand models which control for: whether the system is rank three; whether labour force variables are important determinants of demand; and whether certain household characteristics variables are important determinants of demand. These models are estimated using ML and GMM methods. Whether explanatory variables are non-exogenous is tested, as are the different model specifications, under the different estimation régimes. Also, the relative predictive power of the alternative model specifications is assessed under the ML and GMM estimation régimes.

Evidence is found of non-exogeneity of some explanatory variables. This supports previous research. It is also found that only labour force and age of head of household appear to be important additional determinants of demand. This picture emerges irrespective of estimation method. On the other hand, the predictive power of the various models is poorer when the method of estimation is GMM. This supports the results obtained with Canadian data in Nicol (1995). One important new result which is obtained, owing to the richness of the US price data now available, is that errors in price variables can have large effects on estimation results. The errors in price variables referred to are of the kind one has when households are assumed to face identical prices across the country, when they do not. This result has implications for most of the recent applied demand research using microdata, as this kind of error in price variables is usually present.

The remainder of the paper is structured as follows. In Section 2, model specification is discussed. Some relevant empirical literature is also discussed, which gives direction to the initial model specification.

The data used are discussed in detail in Section 3. Section 4 gives details of the estimated models, hypothesis tests conducted, and indicates comparisons of the predictive performance of the various model parameterisations. Section 5 summarises and concludes.

2 Model Specification

The modelling approach in this paper closely follows Nicol (1995). This has the advantage that some direct comparisons of results can be made of outcomes with Canadian versus US data. Consequently, labour force, household characteristics and other goods which the commodities in the demand system might not be separable from are introduced as "conditioning goods". This is a convenient way to treat many specification issues, since the precise formulation for modelling these aspects does not need to be specified. Whether these variables are important can then be tested using exclusion restrictions on the conditioning goods.

Suppose "goods" over which consumers make decisions can be partitioned into four types. Goods of direct interest, denoted q, and their prices, p; labour force variables, ℓ and their prices, w; other conditioning goods, g and their prices r; and demographic or household characteristics variables, z. If preferences can be represented by the utility function, $U[q, \ell, g, z]$, the conditional cost function is defined as $c[p, \ell, g, z, u] = \min_q[p \cdot q | U(q, \ell, g, z) = u]$. The properties of these functions are discussed in Pollak (1969) and Browning (1983). All variables in q, ℓ, g and z can be vectors. The conditional, compensated demand functions for q are the derivatives of this cost function with respect to p and can be denoted $q_i = f_i[p, \ell, g, z, y]$, where y is total expenditure on the n goods, $q = [q_1, \ldots, q_n]^T$.

The parameterisation used to represent $c[p, \ell, g, z, u]$ is a generalisation of the price independent generalised logarithmic (PIGLOG) model of Muellbauer (1976),

$$\ln c[p, \ell, g, z, u] = \ln a(p, \ell, g, z) + \frac{b(p, \ell, g, z)}{[f(u) - g(p, \ell, g, z)]} \tag{1}$$

The indirect utility function for this model can be written

$$\ln V[p, y; \ell, g, z] = f^{-1}\{\frac{b(p, \ell, g, z)}{[\ln y - a(p, \ell, g, z)]} + g(p, \ell, g, z)\} \tag{2}$$

A model similar to this has been shown to yield rank three, quadratic logarithmic budget-share demand systems with the general form

$$w_i = a_i'(p, \ell, g, z) + \frac{b_i'(p, \ell, g, z)}{b(p, \ell, g, z)}[\ln(y/a[p])] + \frac{g_i'(p, \ell, g, z)}{g(p, \ell, g, z)}[\ln(y/a[p])]^2 \tag{3}$$

by Banks, Blundell and Lewbel (1994). They used a restricted version of this model to estimate a demand system for Britain with FES data. Fry and Pashardes (1992) used an alternative parameterisation, also with British data, and showed that their variant of the model dominated some popular functional forms which are nested within it, such as the Almost Ideal Demand System (AIDS) of Deaton and Muellbauer (1980). The specification in (3) is more general than that in Banks, Blundell and Lewbel (1994), or in Fry and Pashardes (1992), however. Neither directly included conditioning goods or labour force variables. Browning and Meghir (1991), on the other hand, included labour force effects, but did not estimate a rank three system, and did not include any additional, conditioning goods.

The model, (3), encompasses a variety of effects, all of which have been found to be important determinants of demand on their own. Whether these effects should enter the model simultaneously, or whether some are capturing more than one influence must therefore be explored. Given a flexible parameterisation for $a(p, \ell, g, z), b(p, \ell, g, z)$ and $g(p, \ell, g, z)$, this can be analysed. The impact of these separate effects on estimation of by-products of demand estimation (such as prediction, elasticities and equivalence scales) can also be explored.

In Nicol (1995), the functions $a(p, \ell, g, z), b(p, \ell, g, z)$ and $g(p, \ell, g, z)$ were partly chosen due to limited variability in the Canadian price data. This is less important here, since the US price data exhibit much more variability. This will be discussed in the next section. However, to facilitate direct comparisons between the Canadian and US results, the same functions for $a(p, \ell, g, z), b(p, \ell, g, z)$ and $g(p, \ell, g, z)$ are used. These are as follows

$$\ln a(p, \ell, g, z) = \alpha_0(z) + \sum_i \alpha_i(\ell, g, z) \ln p_i + \frac{1}{2} \sum_i \sum_j \gamma_{ij} \ln p_i \ln p_j \tag{4}$$

$$b(p, \ell, g, z) = \beta_0(z) \prod_i p_i^{\beta_i(\ell, g, z)} \tag{5}$$

$$g(p, \ell, g, z) = b(p, \ell, g, z) \cdot \lambda(p, \ell, g, z) \tag{6}$$

$$\lambda(p, \ell, g, z) = \lambda_0(z) + \sum_i \lambda_i(\ell, g, z) \ln p_i \tag{7}$$

where adding up requires that $\sum_i \alpha_i(\ell, g, z) = 1, \sum_i \gamma_{ij} = 0, \sum_i \beta_i(\ell, g, z) = 0$; homogeneity requires that $\sum_j \gamma_{ij} = 0$; and symmetry of substitution effects that $\gamma_{ij} = \gamma_{ji}, \forall i \neq j$. To complete the above specifications requires functional forms for $\alpha_0(z), \alpha_i(\ell, g, z), \beta_0(z), \beta_i(\ell, g, z), \lambda_0(z)$ and $\lambda_i(\ell, g, z)$. Again, the specifications used in Nicol (1995) were employed, which makes it possible to test: exact aggregation; separability of commodity demands from labour supply; separability of commodity demands from other goods; and the importance of household characteristics effects. To accommodate specification of a rank three system, (7) is set equal to

$$\lambda(p, \ell, g, z) = \lambda_0 + \sum_i \lambda_i \ln p_i \tag{8}$$

where $\sum_i \lambda_i = 0$ to satisfy adding up. To introduce the influences of ℓ, g and z into (4)–(6), $\ln a(p, \ell, g, z)$ and $\ln b(p, \ell, g, z)$ are defined:

$$\ln a(p, \ell, g, z) = \alpha_0 + \sum_i \sum_k [\alpha_{ik} v_k] \ln p_i + \frac{1}{2} \sum_i \sum_j \gamma_{ij} \ln p_i \ln p_j \tag{9}$$

$$\ln b(p, \ell, g, z) = \beta_0 + \sum_i \sum_k [\beta_{i0} + \beta_{ik} v_k] \ln p_i \tag{10}$$

where the vector $v = [v_1, \ldots, v_K]^T$ is used to represent ℓ, g and z, for notational convenience. The influences of ℓ, g and z are therefore confined to $\alpha_i(\ell, g, z)$ and $\beta_i(\ell, g, z)$.

Given the above parameterisations for (4)–(7), the following budget-share system can be obtained:

$$\begin{aligned} w_i \quad = \quad & \sum_k \alpha_{ik} v_k + \sum_j \gamma_{ij} \ln p_j + [\beta_{i0} + \sum_k \beta_{ik} v_k][\ln(y/a[p, v])] + \\ & \{\lambda_i + [\beta_{i0} + \sum_k \beta_{ik} v_k][\lambda_0 + \sum_i \lambda_i \ln p_i]\}[\ln(y/a[p, v])]^2 + \epsilon_i \end{aligned} \tag{11}$$

The random term, ϵ_i, denotes a stochastic disturbance such that $[\epsilon_1, \ldots, \epsilon_n]^T \sim \mathrm{N}(0, \Omega)$. The covariance matrix of ϵ is singular, so only $n - 1$ equations of the system need be estimated, the parameters of the n'th can be recovered by the adding-up conditions. Empirical considerations relating to this stochastic specification will be discussed in Section 4.

Many hypotheses can be tested using (11), depending on the restrictions imposed on $\alpha_i(\ell, g, z), \beta_i(\ell, g, z)$ and $\lambda(p, \ell, g, z)$. Restricting the model to a rank two system requires that $\lambda_0 = \lambda_i = 0$, for all i. Separability of commodity demands from labour supply requires that the parameters in $\alpha_i(\ell, g, z)$ and $\beta_i(\ell, g, z)$ on labour force variables be zero, and so on. Also, the interaction of household characteristics variables (in the v vector) with $\ln(y/a[p, v])$ yields a non-exactly aggregable demand system. Six model parameterisations were estimated, and a variety of hypothesis tests carried out. The details of these will be discussed in Section 4.

3 Data

The expenditure data for this study are drawn from the 1980–81, 1982–83, and the annual, 1984–1992 Interview Survey Public-Use Tapes of the CES for the United States. The procedure for collecting data from households in the CES samples was as follows. Each sample was split into three monthly rotation groups. Households in a rotation group were then interviewed in the "same" month of each quarter for five consecutive quarters, reporting their spending patterns for the preceding quarter. This results in one rotation group reporting quarterly expenditures every month. Households can therefore be matched to monthly price data, since it is known when a household reports, and for which quarter. In contrast, in the Canadian case, households report expenditures for the whole year, so this limits the price information which can be used when working with the FAMEX.

At any time, there are approximately 5000 households in the Interview Survey. Pooling data for the eleven indicated surveys permits construction of large, *complete samples* of households groups, even when the groups are homogeneous. Furthermore, given such large *complete samples*, only a proportion of them need be used for exploratory research, leaving similar data available for subsequent estimation of a "preferred" model. Therefore, out of the total households of various types (discussed below), only twenty-five per cent from the *complete samples* were used in this study. These data sets will be referred to as *complete sub-samples*, for reasons which will become clear below. What is learned using the *complete sub-samples* can then be applied in further research involving the remainder of the *complete samples*.

Family size and housing tenure have been found to be of such importance in other demand studies (for example, Barnes and Gillingham, 1984, and Nicol, 1989), as to merit stratifying households into groups based on these variables. Household types were therefore classified according to four different family sizes: married couples without children; married couples with one child; with two children; and with more than two children. Also, three types of housing tenure were used to further classify the households: renter households; home-owners with mortgages; and home owners without mortgages. For all twelve household types, only those with age of head 18–65 and no self-employed members were included in the *complete samples*.

The next step in creating the data sets to be used is the selection of expenditure categories of interest. This choice is governed by a number of considerations. The CES surveys disaggregate expenditures into various categories, so these categories are the minimum level of disaggregation one can work with. However, the kinds of categories in the surveys are similar to those used in surveys in other countries (such as Canada and the United Kingdom). Consequently, the expenditure categories used in this study can be chosen to be as close as possible to those in other studies. This permits some comparisons of results to be made.

The choice of expenditure categories is also dictated by the availability of price data with which to match the expenditures. In addition, the more categories included in a demand system, the greater the number of parameters which have to be estimated. Such estimation is difficult in a nonlinear setting. Large systems can be made smaller by aggregating goods. However, inappropriate aggregation of expenditures can lead to misleading inferences (Nicol, 1991, provides some evidence on this in a homogeneity and symmetry testing context). Consequently, a small, disaggregated demand system is preferable from this perspective. There is then the danger of excluding non-separable goods from such a system. This is not a problem in the present study, however, since "other goods" effects are to be captured by the introduction of conditioning goods.

Given the above considerations, the expenditure categories included in the direct demand system estimated in this paper were food, alcoholic beverages and clothing. All other expenditures were dealt with as an aggregate conditioning good. The categories included in this aggregate conditioning good were: housing; shelter; utilities; household furnishings and operation; transportation (public and private); private transportation; motor fuel; public transportation; medical care; entertainment; personal care. Complete details of these expenditure categories are contained in the Interview Survey Public-Use Tape Documentation.

As the model specification discussion in Section 2 indicated, labour force participation effects were to be included in the demand equations. The CES data contain information on the labour force participation status of adult household members. Consequently, these effects were introduced as labour force participation dummy variables. One variable was included for each of the adult male and female household members. In addition, these dummy variables were interacted with other variables on the right hand side of the estimating equations, as indicated in equation (11).

The discussion above indicates that the households included in the *complete samples* were homogeneous, in terms of the characteristics exhibited. Consequently, there is limited scope for additional household characteristics effects. The nature of the price data (to be discussed below) reflects regional price differences, so there is not much to be gained by inclusion of regional effect variables. These would be highly collinear with the price data, and there would be great difficulty in empirically identifying all the parameters included. One variable which is likely to capture a significant amount of information about the households' characteristics is the age of the head. Also, variables which have been found to be important in other research and which are present in the CES as household characteristics variables are: tobacco consumption by the household; and vehicle ownership. These last two effects were introduced as dummy variables.

Given the *complete samples* of households from the CES, the next step is to match these households to the price vectors which they faced for the goods included in the demand system. There are several other variables in the CES which influence how this matching is done. These are the variables giving household location information. The relevant variables are: region of residence; population size in area of residence; and state of residence. There are four regions of residence: Northeast; Midwest; South; and West. Also, there are five population sizes. These change slightly over the 11 CES data sets used. Also, from 1980–85, the following states were *not* covered by the CES: Delaware, Idaho, Nevada, New Hampshire, North Dakota, Oklahoma, South Dakota, Vermont and Wyoming. From 1986–92, the following states were *not* covered: Montana, Nevada, North Dakota, Rhode Island, South Dakota, Vermont and Wyoming.

For certain households, some or all of the above variables were suppressed, in the interests of confidentiality. However, using these three variables when recorded, it was possible to identify the city in which a household lived. In surveys prior to 1985, one could determine city of residence for 27 US cities by cross-tabulating on the three variables – region, population size and state. These cities are listed in Appendix Table A1. In surveys after and including 1986, however, the population size variables were suppressed for all households in the western states. These states were: Arizona, California, Colorado, Oregon, Utah and Washington. In that case, it was still possible to identify city of residence of households living in twenty cities in the Northeast, Midwest and South, and listed in Table A1. Also, state of residence could be identified for households in West states, and their city of residence inferred as being one of a small sub-set of cities within each of these states.

The change in reporting of household location variables in 1986 had implications for the way that price data could be constructed. For households in cities in the Northeast, Midwest and South, city of residence could be determined for the whole period, 1980–1992. These households could therefore be assigned city prices. On the other hand, for households in the West after 1986, only state of residence could be determined exactly, so these households had to be matched to state prices. This resulted in the introduction of errors in price variables for households in the West after 1986. These types of errors are commonly seen in other data sets (for example, Canadian FAMEX-based data sets, and United Kingdom FES-based data sets). Consequently, the effects of these errors in price variables can be assessed by comparing results based on different *reduced sub-samples* taken from the *complete sub-samples*.

To proceed with the above kind of assessment two kinds of *reduced sub-samples* were constructed from the *complete sub-samples*. The *complete sub-samples* were also used, of course, as part of the assessment, these sub-samples containing households with errors in the price variables. The first set of *reduced sub-*

samples was for households surveyed in all four regions during the period from 1980–85 (referred to as RS1). RS1 therefore covers all regions, but only for a limited period. The second set of *reduced sub-samples* (referred to as RS2) covered households in three regions, the Northeast, Midwest and South, from 1980–92. RS2 therefore covers only part of the US, but for all the sample period.

The household types, MOR0–MOR2, were focused on for this study. Of these, there were, respectively: 758, 565 and 642 observations in RS1; 1345, 959 and 1160 observations in RS2; and 1911, 1338 and 1681 observations in the *complete sub-samples*. RS1 and RS2 were therefore free of the errors in price variables, while the *complete sub-samples* (and the *complete samples*) were not. This permits an analysis of the sensitivity of the results to these kinds of errors in variables, which are very common in studies of this type.

Each of the above three ways of analysing the MOR0–MOR2 data have advantages and shortcomings. The RS1 sub-samples are smaller, and have less variability in prices. But they cover all four regions. The RS2 sub-samples have more variability in prices, but only cover three regions. The *complete sub-samples* are largest and cover all four regions, but the price data contain errors, since some households are assigned state level, not city level prices. However, some insights can be gained into how such estimation results can be affected by the presence of these kinds of errors in price variables. This has not been possible in previous research.

The CES indicate in which month a household is interviewed. In principle, a household could thus be assigned a price vector which would reflect when they made their expenditures. With twelve monthly price observations per year in thirteen years for up to twenty-seven cities, price data would contain a lot of information. Price effects could then be determined precisely. The difficulty is that the CPI data for the US, as for anywhere else, reflect how prices change over time for specific cities. However, city CPI's do not indicate whether prices in New York are higher than in Los Angeles for the same good, such as "food". Fortunately, there is available an extensive database of price indices for six categories, each year, and for many cities. These indices are published by the *American Chamber of Commerce Researchers Association* (ACCRA). The expenditure categories covered are: grocery items, housing, utilities, transportation, health care and miscellaneous goods and services. A representative sample of these data for the cities of interest in this study are presented in Table A1. Using these data in conjunction with city level CPI data, inter-city price indices were constructed for the cities of interest. For example, Table A2 indicates the CPI data for Philadelphia, which are used with the data in Table A1 to yield the data in Table A3. These latter data reflect monthly prices for food, alcohol and clothing for Philadelphia from 1980–92, relative to a US city average with a base of November, 1989. Such prices were constructed for all the twenty-seven cities in Table A1 from 1980–85, and for the reduced (twenty) set of cities from 1986–92. In addition, similar ACCRA data were used in conjunction with CPI data for different city sizes in the West, to construct state level data for 1986–92, for the West states. A representative sample of these is given for Utah, in Table A4. These city and state level price data were then matched to the respective households in the *complete samples*. From these, the *reduced sub-samples*, RS1 and RS2 were then constructed. Further details of the price construction procedure, matching to households, and extraction of actual households from the complete CES are available on request.

4 Estimation and Results

4.1 Exogeneity of Explanatory Variables

It is becoming increasingly common in applied demand studies using microdata for estimation to be conducted using an instrumental variables rather than an ML approach. This is because of concern over purchase infrequency in some bodies of data, which calls into question the exogeneity of total expenditures (y). In addition, if labour force variables are to be included, there is a strong case for controlling for the possible endogeneity of these variables. Also, in the current study, the conditioning of demands on: expenditures outside the demand system; tobacco consumption; and vehicle ownership raises the endogeneity issue too, since such decisions are not independent of allocating expenditures to goods in a food, alcohol, clothing demand sub-system.

Whether instrumental variables estimation is essential in all circumstances of this type, however, should be closely examined. The additional bias introduced in small samples and the increased variability of this estimator are high prices to be paid, if the situation does not require it. In particular, although ML estimation might produce inconsistent estimates if endogeneity of some explanatory variables is not controlled for, these inconsistencies could be partly compensated for by the smaller variances of the ML estimates. Furthermore, estimates based on these parameter estimates such as equivalence scales, elasticities and predictions might not be too inaccurate for the purposes to which they are to be put, whereas instrumental variables estimation based estimates could be much more imprecise. This was illustrated in a recent paper by Nicol and Nakamura (1995), where ML predictions of budget-shares were far superior to corresponding GMM-based predictions. This was shown when using Canadian data, and supporting evidence is also highlighted which appeared in the earlier work of Browning and Meghir (1991), using FES data for the

United Kingdom.

In the paper by Nicol (1995) based on the FAMEX data, a similar result emerged. That is, ML-based prediction was superior to GMM-based prediction, even although the exogeneity of certain explanatory variables was rejected. The current study therefore provides an opportunity to determine whether a similar result holds when using the US CES data.

To assess whether it is necessary to employ GMM estimation, tests of the consistency of ML estimated parameters, $\hat{\theta}$, relative to GMM estimates, $\tilde{\theta}$, were conducted. This was implemented in this paper for each of six model specifications estimated. These model specifications were: a rank three demand system with all variables in v excluded; the same model with age of head; labour force variables; other expenditures; tobacco consumption; and vehicle ownership included[2].

To test whether $\hat{\theta}$ is consistent relative to $\tilde{\theta}$, what Davidson and MacKinnon (1993, p.237) refer to as Durbin-Wu-Hausman (DWH) statistics were used. If the respective estimated covariance matrices of $\hat{\theta}$ and $\tilde{\theta}$ are $\hat{\mathrm{V}}(\hat{\theta})$ and $\tilde{\mathrm{V}}(\tilde{\theta})$, then the statistic $[\tilde{\theta}-\hat{\theta}]^T\{\tilde{\mathrm{V}}(\tilde{\theta})-\hat{\mathrm{V}}(\hat{\theta})\}^{-1}[\tilde{\theta}-\hat{\theta}] \overset{A}{\sim} \chi^2(q)$, where q is the rank of the inverse of the covariance matrix difference, under the null hypothesis that the ML estimator is consistent (Hausman, 1978). Unfortunately, in the present application, this statistic did not prove useful, since the matrix difference forming the inner-product was not positive definite[3].

To circumvent the above problem, and still test for the consistency of $\hat{\theta}$, an alternative test statistic suggested by Davidson and MacKinnon (1993, p. 238–242) was used, on an equation-by-equation basis. This test statistic is based on the artificial regressions

$$w_i - f_i(\hat{\theta}) = X(\hat{\theta})b + M_W X(\hat{\theta})^* c + \text{residuals} \tag{12}$$

where $f_i(\hat{\theta})$ is the nonlinear function in the budget-share equations evaluated at the ML estimate, $X(\hat{\theta})$ is a matrix of derivatives of $f_i(\theta)$ evaluated at the ML estimate, and $X(\hat{\theta})^*$ is a sub-matrix of $X(\hat{\theta})$, containing the columns of $X(\hat{\beta})$ which are not in the span of W asymptotically. The matrix W is the matrix of instrumental variables, and $M_W = [I - P_W] = [I - W(W^TW)^{-1}W^T]$. The instruments used are described in Appendix B. Some descriptive statistics on all the variables used are also contained in this Appendix.

A test of the null that c in (12) is zero is an asymptotically valid F test of the null hypothesis that $\hat{\theta}$ is consistent. This statistic has $k^*, [N-k-k^*]$ degrees of freedom, where k is the row dimension of θ, N is the sample size, and k^* is the number of explanatory variables in the model which are not exogenous.

[2]The results of testing these competing models is discussed in the next sub-section.

[3]This is a common problem with this test statistic in finite samples.

The results of the tests of consistency of $\hat{\theta}$, or more loosely speaking, exogeneity of X, are presented in Tables 1–3 for the two equations estimated, given that one equation is dropped because of the adding-up conditions. These tables are, respectively, based on estimation for: the *complete sub-samples*; the RS1 sub-samples; and the RS2 sub-samples.

Table 1 indicates strong evidence against the null hypothesis for MOR0 and MOR2, but less strong for MOR1. However, recall that these results are based on the data with the errors in the price variables. Table 2 does not, however, provide strong evidence against the null at all. The sample sizes these results are bases on are smaller than those for Table 1, but there is still a high degree of price variability, although the time-period covered is shorter. Referring next to Table 3, only for MOR2 is there strong evidence against the null. These results are based on the RS2 samples. That is, larger than the RS1 samples, smaller than the *complete sub-samples*, but with none of the errors in price variables. It would therefore appear that caution should be exercised when interpreting the results of DWH-type tests with data of this kind.

It is also interesting to note that *less* restrictive models tend to reject the null relatively more frequently, which indicates caution should be employed by researchers using more simple models, when they find non-rejections. In addition, one could argue that the results here are more persuasive than those based on Canadian data, since those data contain quite large errors in the price variables. Owing to the geographic information provided in the FAMEX, and the available Canadian price data, one has to assign prices based on an Ottawa-Toronto average to all residents of Ontario, for example. Alternatively, all Québec residents have to be assigned indices for Montréal. The errors in price variables induced by this approach are certainly at least of the same order of magnitude as those made in the present study by assigning all residents of states in the West to price variables for their state of residence.

In any event, there are researchers who would still not be comfortable relying on ML-estimates results. Consequently, in what follows, the results presented are based on both ML and GMM estimation of the models of interest, using the MOR0–MOR2 *complete sub-samples*, and the *reduced sub-samples*, RS1 and RS2. This makes it possible to compare test results in the next sub-section based on the different estimation methods, and to compare the alternative predictive performance of the two methods also.

4.2 Model Specification Tests

In previous work using Canadian microdata conducted by the author, in general, it has been found that homogeneity and symmetry restrictions were supported by the data for the types of model being estimated here. Furthermore, these restrictions seem to be supported on the whole when using microdata from a variety of countries. Consequently, these restrictions are imposed in this study.

Model specification tests involving elements of the vector, v were conducted based on ML and GMM estimation[4] of the six models described in the preceding sub-section. Likelihood Ratio tests based on ML estimation are given in Table 4. To summarise the contents of this table, only labour force and age of head variables appear to be important additional determinants of demand for the MOR2 and MOR0–MOR2 households respectively. This is close to the finding in Nicol (1995), where the same variables were found to be important determinants of demand, although in that case, the labour force variables were important for MOR0–MOR2. Also, in Nicol (1995) the other expenditures conditioning good was found to be important, but this was not the case here. Finally, it was also found that it was not necessary to model the demands as a rank three system. This last finding is not too surprising. Lewbel (1989) found that budget-shares linear in $\ln(y/a[p])$ (rank two demand systems) were adequate to fit United Kingdom FES data, when the households in the samples were homogeneous. The households in the samples used here were constructed to be as homogeneous as possible. This result also supports a similar finding in Nicol (1995).

Hypothesis tests based on the GMM estimates were conducted under the Wald Principle, since this was more convenient given the estimation method used. Exclusion restrictions were tested based on the least restrictive model, for omission of: age of head variables; labour force participation variables; other expenditures variables; tobacco consumption variables; and vehicle ownership variables. These test results are reported in Table 5.

The GMM test results showed some conflict with the ML test results. While labour force variables were again found to be important, this was for the MOR0 group rather than MOR2. Also, age of head variables were important, but only for MOR1 and MOR2, and the other expenditures conditioning good was important for MOR0 and MOR1. Again, a rank two system seemed to be sufficient on the whole, but vehicle ownership now seemed to be important for MOR0. Neither the ML or GMM estimation-based test results, however, seemed to be greatly affected by the errors in price variables, in contrast to the results in the preceding sub-section.

In conclusion, there does seem to be evidence that labour force and age of head variables are important, irrespective of the method of estimation. Some finer points are obscured by the ability of tests to discriminate between competing models when many parameters are being estimated (35 for the unrestricted

[4]ML estimation was conducted using SHAZAM, Version 7.0; GMM estimation was conducted using TSP, Version 4.3.

model), in the context of a highly nonlinear model. One important point which seems to emerge fairly clearly, however, is that it does not seem to be necessary to model demands as rank three systems, with such homogeneous household group samples. Such a simplification greatly reduces the complexity of the estimation problem, and indicates an appropriate way to parameterise the demand model, without leaving out something of importance. This is a finding of some significance for future work and is all the more useful since it seems to be quite generally true (that is, when using British, Canadian and US data).

4.3 Predictive Power Under Alternative Estimation Conditions

Even although the exogeneity of certain explanatory variables was rejected in Nicol (1995), it was found that prediction of budget-shares was much poorer when based on GMM estimated parameters. In this sub-section, this issue is explored in the context of the CES data.

To explore predictive power, goodness-of-fit, pseudo χ^2 statistics were calculated. The method is to take the fitted values of budget-shares under the alternative estimation regimes, and compare the performance of the predictions for different household sub-groups. The sub-groups were based on labour force participation status of the adults in the households. There were four strata of labour force participation in any given household. Both working outside the home, one working outside the home; and both not working outside the home.

Denote ML predictions by $\hat{w}_i^{ml}$ and GMM predictions by $\hat{w}_i^{iv}$, and actual shares by w_i^A. Then the deviations of actual from predicted shares are d_i^{ml} and d_i^{iv} for ML and GMM respectively. The sampling distributions of the means of these deviation vectors are then normally distributed, by the Central Limit Theorem. There are three shares, but only two independent pieces of information. These sample means can be converted to standard normals, and χ^2 statistics calculated. That is, we have $\sum_i^3(\bar{z}_i^{ml})^2 \sim \chi^2(2)$ and $\sum_i^3(\bar{z}_i^{iv})^2 \sim \chi^2(2)$ also. Here, $\bar{z}_i^{ml} = [1/N]\sum_h^N \sqrt{N} d_i^{ml}/\hat{\sigma}_{ml}$, where N is the number of households in a given labour force participation stratum, h indexes households and $\hat{\sigma}_{ml}$ is the sample standard deviation of the vector of deviations, d_i^{ml}. Analogous definitions can be made for the corresponding GMM case.

The above χ^2 statistics were calculated for each of the six models and estimation régime, across all four labour force participation strata. In addition, each of the *complete sub-sample* and *reduced sub-sample* data groups were used, so that the role of errors in the price variables discussed earlier could be analysed in the context of predictive power. The χ^2 statistics are an indication of the goodness-of-fit of the respective estimates, and are presented in Tables 6–8 for the *complete sub-samples*; the RS1 *reduced sub-samples*; and the RS2 *reduced sub-samples*. From this, it can be seen that predictive power is almost always much better when estimation is by ML. The only exceptions to this are in cases where labour force strata contain too few observations to obtain reliable χ^2 statistics. This confirms the earlier discussion that the GMM estimates could possibly have larger variability than the ML estimates, even though the latter could exhibit inconsistency.

Out of a total of eighteen possible comparisons, predictions based on ML estimates are superior in fifteen cases. Further, there is only one cell which has a lower χ^2 statistic for ML than GMM in two out of the three other cases. One of the reasons for using GMM was the possible endogeneity of labour force participation variables. One might therefore expect that at least for models with labour force participation variables present, the GMM might perform fairly well. This would be for Models 4–6. It can be seen, however, that the GMM predictions are poorer than the ML predictions in all of these cases.

These results indicate that there are some dimensions at least where it matters a great deal whether estimation is by ML or by GMM. Furthermore, this supports the earlier results in Nicol (1995). It will be interesting to explore whether this kind of finding holds for other estimates implied by demand parameters, such as equivalence scales, elasticities and measures of inequality. These are all uses to which estimated demand parameters are put. Research on these other issues is continuing.

5 Summary and Conclusions

In this paper, a general model of demand was estimated which was a rank three system. The model was also specified to control for the possibility of non-separability of labour force participation effects, non-separability of other goods, not included directly in the demand system, and the influence of household characteristics effects.

Recent research has focused on the possible need to control for non-exogeneity of certain explanatory variables. As a result, this aspect was investigated by conducting tests for the consistency of ML estimates. These tests indicated that non-exogeneity could not be ruled out. However, these results were sensitive to the presence of errors in the price variables used. As a result, comparative results were presented for various model specification hypothesis tests using ML and GMM estimated parameters. All of these estimates and tests used US CES data for the years 1980–1992. This paper is unique in this respect. Past studies have not used these microdata in this way, owing to difficulties in obtaining the requisite price data to assign to households living in different regions, in different time periods.

The foregoing difficulty with the price data was overcome by constructing price data based on US CPI data, and ACCRA data, the latter reflecting price differences across different US cities at a point in time for various expenditure categories. These price data were constructed for twenty-seven US cities, and it was possible to determine from the CES files which of these cities most of the households in the samples used resided in. This resulted in very rich data sets, covering homogeneous households groups, this richness being exhibited both in terms of price variability and household characteristics. In addition, because it was possible to assign only state level price data to some households (owing to the nature of the available data), it was possible to analyse the effects of errors in price variables on the estimation result, as mentioned above with respect to the exogeneity test results.

It was found that labour force and age of head variables were important additional determinants of demand, compared to the basic model. However, a rank two demand system was found to be adequate, rather than the rank three system specified initially. These results were true irrespective of whether estimation was by ML or GMM. These findings support the earlier results in Nicol (1995), using the FAMEX data.

Other test results were more ambiguous, partly because of the presence of errors in the price variables for some households, but also because of the alternative estimation methods yielding competing implications. In future work, since the homogeneous household groups only seem to require a rank two system, it is intended to employ this more simple specification. This should enable more precise hypothesis tests to be conducted, as many standard errors of estimated parameters are quite large for the unrestricted model, which contains thirty-five parameters, when in its rank three form.

A comparison of predictive power under the alternative estimation régimes of ML and GMM was then conducted, and this indicated that the former method produced much better predictions. Thus, while there seems to be some evidence that GMM estimation is required on the basis of non-exogeneity of some explanatory variables, this does not seem to make a difference for the purposes of hypothesis testing, or for prediction. Furthermore, the exogeneity test results appear to have been influenced by the errors in price variables for some observations. It is thus not clear that the GMM procedure is absolutely necessary.

The work in this paper was intended to provide some insights into the direction of future research. However, it is also complementary to the research in Nicol (1995), and confirms a number of the results in that paper. The data sets used were deliberately chosen to ensure that a detailed analysis could be performed on the importance of certain model specification issues. At the same time, sufficient additional data remains to analyse the more refined models indicated by the present study. This research is continuing.

References

American Chamber of Commerce Researchers Association, Inter-City Price Indices for the United States. Various years.

Banks, J., R. Blundell and A. Lewbel (1994), "Quadratric Engel Curves, Indirect Tax Reform and Welfare Measurement." Discussion Paper No. 94–04, University College London.

Barnes, R. and R. Gillingham (1984), "Demographic Effects in Demand Analysis: Estimation of the Quadratic Expenditure System." *Review of Economics and Statistics*, **66**, 591–601.

Blundell, R., P. Pashardes and G. Weber (1993), "What Do We Learn About Consumer Demand Patterns From Micro Data." *American Economic Review*, **83**, 570–597.

Browning, M. (1983), "Necessary and Sufficient Conditions for Conditional Cost Functions." *Econometrica*, **51**, 851–856.

Browning, M. and C. Meghir (1991), "The Effects of Male and Female Labor Supply on Commodity Demands." *Econometrica*, **59**, 925–951.

Bureau of Labor Statistics, Washington DC. Interview Survey Public-Use Tapes, Consumer Expenditure Surveys, 1980–1992.

Dagenais, M. (1994), "Parameter Estimation in Regression Models With Errors in Variables and Autocorrelated Disturbances." *Journal of Econometrics* **64**, 145–163.

Davidson, R. and J.G. MacKinnon (1993) *Estimation and Inference in Econometrics*, Oxford University Press.

Deaton, A. and J. Muellbauer (1980), "An Almost Ideal Demand System." *American Economic Review*, **70**, 312–326.

Dickens, R., V. Fry and P. Pashardes (1993), "Non-linearities and Equivalence Scales." *Economic Journal*, **103**, 359–368.

Fry, V. and P. Pashardes (1992), "An Almost Ideal Quadratic Logarithmic Demand System for the Analysis of Micro Data." Discussion Paper No. 25, City University of London.

Hausman, J.A. (1978) "Specification Tests in Econometrics". *Econometrica* **46**, 1251–1272.

Jorgenson, D.W, L.J. Lau and T.M. Stoker (1982), "The Transcendental Logarithmic Model of Aggregate Consumer Behavior". In *Advances in Econometrics* (Vol. 1), R.L. Basmann and G.F. Rhodes, eds, JAI Press.

Kaiser, H. (1993), "Testing for Separability Between Commodity Demand and Labour Supply in West Germany." **18**, 21–56.

Lewbel, A. (1989), "Household Equivalence Scales and Welfare Comparisons." *Journal of Public Economics*, **39**, 377–391.

Muellbauer, J. (1976), "Community Preferences and the Representative Consumer." *Econometrica*, **44**, 525–543.

Nicol, C.J. (1989), "Testing a Theory of Exact Aggregation." *Journal of Business and Economic Statistics*, **7**, 259–265.

Nicol, C.J. (1991), "The Effect of Expenditure Aggregation on Hypothesis Tests in Consumer Demand Systems." *International Economic Review*, **32**, 405–416.

NIcol, C.J. (1995) "Model Specification and Estimation Effects in Applied Demand Analysis Using Microdata". Department of Economics Working Paper No. 52, University of Regina.

Nicol, C.J. and A. Nakamura (1995), "The Effects of Labor Supply on Commodity Demands: Some Canadian Evidence." Department of Economics Working Paper No. 53, University of Regina.

Pollak, R. (1969), "Conditional Demand Functions and the Implications of Separability." *Quarterly Journal of Economics*, **83**, 70–78.