The Effects of Labor Supply on Commodity Demands: Some Canadian Evidence†

by

Christopher J. Nicol	and	Alice Nakamura
Department of Economics		Faculty of Business
University of Regina		University of Alberta
Regina, Saskatchewan, S4S 0A2		Edmonton, Alberta, T6G 2R6
CANADA		CANADA
306-585-4182		403-492-5824

February 9, 1997

†Nicol's and Nakamura's research were supported by separate grants from the Social Sciences and Humanities Research Council of Canada. Funding for the research in this paper was also provided by a grant from the Edmonton Social Planning Council as part of a larger project funded by Health and Welfare Canada, and by support for Nicol from the University of Regina President's Research Fund. We are grateful to Adolf Buse and Marcel Dagenais for comments on an earlier version of this paper. All errors are the responsibility of the authors.

Abstract

This paper confirms the finding of Browning and Meghir (1991) (who use United Kingdom microdata), that commodity demands are *not* separable from labor supply variables, using Canadian microdata. They also present evidence that omitting these labor supply variables from demand models seriously biases estimated price, expenditure and child status effects.

However, closer examination of this latter issue indicates that the magnitude of biases is *greater* when using instrumental variables to capture the labor supply effects, than merley including the labor supply variables themselves.

Keywords:

consumer demand; labor supply; separability; microdata.

JEL Classification:

C13; C31; D12; J22.

1 Introduction

In an innovative and influential paper, Browning and Meghir (1991) present estimation results for a conditional demand system that is an extension of Pollak's (1969, 1971) conditional demand approach. The conditioning "goods" are labor supply variables: the labor force participation status (employed versus not employed) and the usual hours of work per week of the husband and the wife. Browning and Meghir (hereafter, BM) show how a test for the separability of commodity demands from these labor supply variables can be carried out, in a relatively simple manner. Using data from the United Kingdom Family Expenditure Survey (FES) for 1979 through 1984, they find that separability is rejected. Furthermore, they present evidence which they interpret as establishing that ignoring the effects of labor supply variables in estimating household demand models can lead to substantial biases in estimated price and expenditure effects, as well as in estimated child status effects.

In our study, the BM rejection of separability of commodity demands from labor supply variables is replicated using data from the 1982, 1984 and 1986 Canadian Family Expenditure Surveys (FAMEX). This study also explores the consequences of instrumenting the labor supply and total expenditure variables in the conditional demand model, as BM do. We find the apparent magnitudes of the biases due to omitting the labor supply variables are greatly affected by whether the labor supply variables are directly entered, or instrumented.

Section 2 outlines the modelling approach used in the BM study, and Section 3 discusses their instrumental variables (IV) treatment of the labor supply and total expenditure variables. Section 4 describes our data and variables. The estimated models are given in Section 5, and our separability results are presented in Secion 6. Conflicting evidence concerning the severity of coefficient bias problems is considered in Section 7. Section 8 concludes. Summary statistics and complete sets of our estimated coefficients are shown in the Appendix.

2 The Browning-Meghir Study

2.1 Formulation

Three categories of variables are defined in the BM study: *goods of interest*, with the quantity and price vectors, q and p respectively; *conditioning goods*, with the quantity and price vectors h

and r respectively; and *demographic variables*, denoted by the vector, a. As BM show, if the goods of interest are weakly separable from the conditioning goods, then the conditional demand system has the form

$$q_i = f_i(p, a, y) \tag{1}$$

where y is the total expenditure on the goods of interest. When the goods of interest are not separable from the conditioning goods, then the conditional demand system has the form

$$q_i = f_i(p, h, a, y) \tag{2}$$

Hence a simple test of weak separability consists of testing whether the demands (q_i) depend on the quantities of the conditioning goods (h) after controlling for the prices of the goods of interest (p), the total expenditure on these goods (y), and the relevant demographic factors (a).

In the empirical portion of the BM study, the categories of goods of interest are Food, Alcohol, Fuel, Clothing, Transportation, Services and Other Goods. As noted in Section 1, the conditioning goods are represented by the weekly hours of work for the husband and for the wife (h_m and h_f respectively), and dummy variables for the employment status of the husband and the wife (d_m set equal to 1 when $h_m > 0$, and d_f set equal to 1 when $h_f > 0$). The demographic variables in the BM study are the numbers of children aged 0–4 and 5–18, the age of the wife, three dummy variables for the birth cohort of the husband (for pre-1931, 1931–40 and 1941–50 respectively), three seasonal dummy variables, and ten regional dummy variables. The total expenditure variable (y) is defined to be the aggregate expenditure on the seven categories of goods if interest, deflated by a household-specific price index.

Using an Almost Ideal parameterisation (see Deaton and Muellbauer, 1980), the typical budget-share equation in the BM study is given by

$$w_i = \alpha_i[h_m, h_f, d_m, d_f, a] + \sum_j \gamma_{ij} \ln p_j + \beta_i[h_m, h_f, d_m, d_f, a] \ln y + \epsilon, \tag{3}$$

where ϵ_i denotes an error term. BM impose both homogeneity and symmetry. That is, they impose the conditions that $\sum_j \gamma_{ij} = 0$ for all i, and $\gamma_{ij} = \gamma_{ji}$ for all $i \neq j$. The Other Goods equation is dropped to accommodate adding up. In the tables in BM, estimates of the α_i's and the γ_{ij}'s are presented under the sub-heading of *intercept coefficients*, while the estimates of the β_i's are presented under the sub-heading *total expenditure coefficients.* We also follow this practice in the tables in this paper.

2.2 Use of Instrumental Variables

BM state that:

> An immediate reaction to a conditional demand system of the form given ... is that it includes variables on the right-hand side that may be endogenous for the budget share equations ... In the demand system we estimate there are five (potentially) endogenous variables. These are: male and female hours of work, male and female participation, and total expenditure ... To identify the parameters in the system we assume that, conditional on total expenditure and the labor variables, asset income and education do not enter the demand system. Given this assumption the model is identified (using these instruments and other functions of them ...)
>
> *Browning and Meghir, 1991, pp. 932–933.*

The complete set of instruments that they use to compute predicted values for the male and female labor supply variables and total expenditure are:

- male and female education, and education squared
- male age and male age squared
- female age squared
- the numbers of children in each of four age groups
- asset income, and asset income squared
- the asset income variables crossed with each of the four child status variables
- the log price variables for the seven categories of goods crossed with the total number of children
- survey year dummy variables
- all of the remaining variables on the right-hand side of the budget-share equations except, of course, the labor supply and total expenditure variables.

The nature of the instrument set chosen in the BM study is typical of the demand literature approach to the possible endogeneity of explanatory variables. However, this methodology raises the possibility of a loss of precision due to multicollinearity, owing to the cross-product terms which are included, and their consequent correlation with included right-hand side variables not considered endogenous.[1] This issue will be discussed in more detail later.

3 Our Data, Demand System and Variables

Our data are drawn from the 1982, 1984 and 1986 FAMEX surveys. Only households that are married couple families with the husband and wife both between the ages of 18 and 65 are included in our estimation data sample. Households with self-employed adults have also been excluded. Table A.I in the Appendix shows the numbers of households in our estimation sample by survey year and child status, while Table A.II shows the numbers of households by the work status of the husband or the wife, and the household child status.

Our demand system is for four categories of household expenditures that are designated as: Food, Clothing, Utilities and Services. Further details of the types of expenditures included in each of the four categories are given in Table I. An Almost Ideal parameterisation is used, as in the BM study. Hence the basic form of our budget-share equations is given by (3).

Our price variables were formed by first using Statistics Canada inter-city retail price indexes for more disaggregate expenditure categories to form regional price indexes, and then aggregating these to form price variables for our aggregate expenditure variables (see Nicol, 1989, for details).

As measures of the amounts of labor supplied by those who were employed, we use variables for the number of full-time and the number of part-time weeks of work in the given years. Our data source provides no hours of work information.

We use three dummy variables to denote husbands aged 40–49, 50–59 and 60 and over respectively. These dummy variables are roughly comparable to the cohort dummy variables in the BM

[1]The endogeneity of explanatory variables also arises frequently in the empirical labor economics literature. The problem there is similar to the kind of endogeneity which is likely in this study. Nakamura and Nakamura (1992) survey this and other issues as they relate to the effects of children on female labor supply. Many of the observations of Nakamura and Nakamura (1992) regarding the endogeneity of explanatory variables in the labour economics literature, and the ways of controlling for that endogeneity, are relevant to our paper.

study. Also, like BM, we include regional dummy variables. These are for the Atlantic Provinces, Québec, the Prairie Provinces, and British Columbia (the omitted category is Ontario).

4 Four Estimated Models

We wish to examine the consequences of including or omitting the labor supply variables. More specifically, we want to test the hypothesis of weak separability of commodity demands from the labor supply variables, and to explore the nature of the omitted variable biases if weak separability is rejected.

The BM test of weak separability and their evidence on omitted variable bias is conditional on the maintained hypothesis that the labor supply and expenditure variables are endogenous and the choice of an instrument set for these variables. With this in mind, a second objective of our study is to gain insight into the potential importance of the presumption and treatment of endogeneity in the BM study. In order to do this, we present estimation results for four model variants, rather than two as BM do. Our estimated models are:

1. full budget-share equations with the labor supply and total expenditure variables included with these entered directly, rather than instrumented, termed the *full:direct* model;

2. budget-share equations with the labor supply variables omitted and the total expenditure variables entered directly, termed the *restricted:direct* model;

3. budget-share equations with the labor supply variables omitted and the total expenditure variables instrumented, termed the *restricted:IV* model;

4. full budget-share equations with the labor supply variables included and with these and the total expenditure variables instrumented, termed the *full:IV* model.

In our terminology, BM present separability test results based on the *full:IV* model, and draw their conclusions about the magnitude of omitted variable coefficient bias problems when the labor supply variables are omitted from comparisons between estimation results for the *full:IV* and the *restricted:IV* models.

The instrument set we have used for our *full:IV* and *restricted:IV* model variants is in accord with the instrument set used by BM. The variables we use as instruments are:

- dummy variables for male and female education at different levels
- male age, and male age squared
- female age squared
- the numbers of children in each of six age groups
- asset income, and asset income squared
- the asset income variables crossed with each of the six child status variables
- the log price variables for the four expenditure categories of goods, crossed with the total number of children
- housing tenure
- survey year dummy variables
- all of the remaining variables on the right-hand side of the budget-share equations except, of course, the labor supply and total expenditure variables.

The coefficient estimates (all multiplied by 100) for the *full:IV* and *full:direct* models are shown in Table A.III in the Appendix, and the coefficient estimates for the *restricted:IV* and *restricted:direct* models are shown in Table A.IV. In both of these tables, t-statistics based on heteroskedasticity-robust standard errors are given in parentheses below the coefficient estimates. Test results and selected groups of the estimated coefficients are discussed in the following sections of the paper.

5 Rejection of Separability

If the demands for the commodities of interest are weakly separable from labor supply, then the coefficient estimates for the labor supply variables should be insignificantly different from zero. But which are the appropriate estimates of the labor supply coefficients–those from the *full:IV* or those from the *full:direct* model?

5.1 *Results with Instrumental Variables*

BM implicitly assume that if the demands for the commodities of interest are not weakly separable from labor supply, then the labor supply variables (and total expenditure) must be jointly and simultaneously determined along with the commodity demands in the current time period. If this is the case, then coefficient estimates for directly included labor supply variables may be biased estimates of the true labor supply effects and may yield misleading results for tests of labor supply effects on the commodity demands. Hence the separability test results that BM show are for the *full:IV* model.

Our labor supply coefficient estimates for the *full:IV* model are shown in columns 1, 3 and 5 of Table II. For each expenditure category, it can be seen that many of the *full:IV* labor supply coefficient estimates are significantly different from zero, at least at an 80 percent confidence level.

System wide separability test results for the *full:IV* model, for the exclusion of all and of subsets of the labor supply variables, are shown in the top half of Table III. These test statistics are based on the Wald principle, and use a heteroskedasticity-robust estimate of the variance-covariance matrix for the estimated parameters. Corresponding test results to those shown in Table III were also computed without correcting the variance-covariance matrix for heteroskedasticity. The pattern of results is identical to what is shown in Table III. The labor supply variables are significant as a group for the *full:IV* model.

The implication of both the Table II and the Table III test results for the *full:IV* model is that the BM rejection of weak separability for commodity demands from labor supply variables is supported.

5.2 *Results with Labor Supply and Expenditure Variables Directly Included*

Of course, the current endogeneity of right hand side variables in no way means that a particular instrumental variables treatment of those variables will eliminate all endogeneity bias problems. Some of the variables in the instrument set may be correlated with the true disturbance term(s). Tests for this will essentially depend on assertions that yet another set of variables are uncorrelated with the true disturbance term(s). What this means is that there is no certain way of proving the validity of the instrumental variables treatment for the labor supply variables in the *full:IV* model, in either this study or the BM one[2]. Nor is it necessarily the case that the IV estimates of the labor supply coefficients in the *full:IV* budget-share equations are less biased in comparison with the labor supply coefficient estimates for the *full:direct* model. Furthermore, as researchers using Monte Carlo experimentation are well aware, the mean-squared error (MSE) for biased estimators can often be *smaller* than the MSE for other, consistent estimators of the same parameters.

One reason for the foregoing type of MSE result is that bias can be larger in small samples for consistent estimators than inconsistent ones. Also, in the present situation, additional "noise" in the instrumental variables increases standard errors. There is the possibility of multicollinearity of constructed regressors with exogenous regressors causing standard errors to be even higher. A comparison of correlations between exogenous right hand side variables and endogenous right hand side variables under two conditions illustrates this. The replacement variables under instrumental variables estimation are much more highly correlated with right hand side exogenous variables than the original, endogenous right hand side variables. Consequently, precision of the *full:IV* estimates is not only reduced for the usual reason associated with IV estimation, but also because the columns "design matrix" for *full:IV* estimation are highly correlated. This is also likely to be the case in the BM study, as they follow the usual practice in this literature for their choice of instruments, as discussed earlier. In our study, we also follow this approach, since we are replicating the BM methodology in the context of Canadian microdata.[3]

[2]In working paper versions of the published BM article, it is clear they had some problems in determining an appropriate set of instruments. That is, tests of the identifiability of the model parameters rejected exclusion of groups of the initial instrument set. However, non-rejection of the null hypothesis that instruments can be excluded can still be problematic. In analytical and Monte Carlo work, Cragg and Donald (1993) found that the actual size of certain identifiability tests can be orders of magnitude smaller than their nominal size. This gives cold comfort in situations where one observes non-rejection of identifiability restrictions: rather than indicating correct model specification, such test results could merely be signalling non-identifiability. Thus, the instrument set *cannot* be excluded; perhaps because they are highly correlated with the variables they are intended to replace, yet *still* correlated with the model's disturbances. In such a case, using the "endogenous" variables is likely to be preferable to using the "instruments", since the latter suffer from the same problem as the former, yet induce larger biases than the variables they are intended to replace (since instrumental variables estimators will, in general, be biased, and have large variances).

[3]As mentioned above, the empirical labor economics literature is also faced with the problem of dealing with endogenous explanatory variables. While looking for other instrumental variables is one way of dealing with a poor

Another possible factor contributing to the imprecision of the *full:IV* and *restricted:IV* estimates is the low correlation between some of the instrumental variables and the endogenous variables they replace. That is, the multiple correlation coefficients for many on the endogenous variables, when regressed on the instrument set, were very low. This indicates that the instrumental variables are not very good proxies for the variables which they are replacing. Unfortunately, little can be done about this since, in the BM study, all potentially appropriate variables were included in the instrument set for this study, given the data sets being used.

These considerations lead us to be interested in comparing coefficient estimates for the labor supply variables and both t-test and system wide test results for the significance of labor supply effects, for the *full:direct* versus the *full:IV* model. There is a second reason for our interest in this comparison. It seems conceivable that labor supply (and also total expenditure) may be adjusted less frequently than demands for specific goods categories within the expenditure bundle. From this perspective, the labor supply variables could affect the demands for the commodities of interest without being jointly determined *in the current period* with the commodity demands. That is, it seems possible that the weak separability of commodity demands from labor supply is false, but that the labor supply variables are predetermined rather than being current endogenous variables.

Our labor supply estimates for the *full:direct* model are shown in columns 2, 4 and 6 of Table II. For each expenditure category, it can be seen that many of the *full:direct* labor supply coefficient estimates are significantly different from zero at least at an 80 percent confidence level. Also, system wide separability test results for the *full:direct* model are shown in the bottom half of Table III. These results imply that the labor supply variables are significant as a group for the *full:direct* model.

Thus, the weak separability of demands for the goods of interest and labor supply is rejected

choice of instruments, the empirical labor economics literature provides examples of other approaches. Rosenzweig and Wolpin (1980) use the exogenous variation of multiple births as an instrument in a life-cycle labor supply model. On the other hand, Rosenzweig and Schultz (1985) use a "fecundity factor" as an element in the estimation of the child-status effects variables included in a model of female labor supply. Since female labor supply is strongly influenced by child-status effects, as documented in the empirical labor economics literature, the approaches of Rosenzweig and Wolpin (1980) and Rosenzweig and Schultz (1985) could be modified to deal with the endogeneity of labor force participation variables in a demand model.

whether or not the labor supply variables (and total expenditure) are treated as being current endogenous variables.

6 Conflicting Evidence on the Severity of the Coefficient Bias Problems

6.1 The BM Evidence on Bias Severity

Having rejected the separability of commodity demands and labor supply, BM explore how omitting the labor supply variables affects the estimated values of coefficients for key variables in the demand model. They compare the coefficient estimates for the *full:IV* model, the *restricted:IV* model, and for a third model that includes the instrumented dummy variables for participation in the work force but not the weekly hours variables. They draw the following general conclusions:

> Estimates of demand that only take account of participation (and not hours) are likely biased. Estimates of demand that ignore labor supply altogether are subject to even more bias.
>
> *Browning and Meghir, 1991, p. 944.*

These conclusions are important, if correct, because most previous demand studies do ignore labor supply altogether. In this section, the nature of the evidence on which these conclusions are based is re-examined.

BM focus on the price and child status coefficient estimates in reaching their bias conclusions. The estimation of price responses has been a central concern in the demand analysis literature. As for the child status coefficients, BM write:

> One area that we conjectured would be subject to biased inference if we ignore labor supply information was the effects of children on demand. The presence of children (and particularly young children) is highly correlated with the participation regime. Given this, we would expect that ignoring the participation variable would cause significant changes in the coefficients on children.
>
> *Browning and Meghir, 1991, p. 942.*

BM focus on the differences in the values for their *full:IV* and *restricted:IV* coefficient estimates for the price and child status variables. The assumption underlying their analysis of these coefficient differences is that the differences are due to bias problems with the *restricted:IV* estimates that are caused by the omission of the labor supply variables from that model.

6.2 Our Full:IV *and* Restricted:IV *Coefficient Estimates*

Our *full:IV* estimates are shown in column 4 of Table IV for the price variable coefficients and in column 4 of Table V for the child status variable coefficients. The corresponding restricted model coefficient estimates are shown in column 3 of Table IV for the price variables and in column 3 of Table V for the child status variables. These are coefficient estimates for our *restricted:IV* model, with the labor supply variables omitted and the total expenditure variables instrumented, which corresponds to the BM restricted model.

BM focus on the differences they find between their *full:IV* and *restricted:IV* models. Looking at our *full:IV* (column 4) and *restricted:IV* (column 3) price coefficient estimates in Table IV, the sign patterns are the same but the magnitudes of the estimated coefficients exhibit large percentage differences. Also, fewer of the *full:IV* price coefficient estimates are found to be significantly different from zero. Looking now at our *full:IV* and *restricted:IV* child status coefficient estimates in Table V, there are differences in the sign patterns as well as the magnitudes, and far fewer of the *full:IV* coefficient estimates are found to be significantly different from zero (only two, even with a .20 critical region).

Like BM, we could interpret the observed differences between our *full:IV* and *restricted:IV* coefficient estimates as being due to omitted variable biases in the *restricted:IV* estimates attributable to the omission of the labor supply variables.

6.3 Our Full:Direct *and* Restricted:Direct *Coefficient Estimates*

Our *full:direct* and *restricted:direct* coefficient estimates are shown in columns 1 and 2 of Table IV for the price variables and in columns 1 and 2 of Table V for the child status variables. We find these two sets of coefficient estimates, for the *full:direct* and *restricted:direct* models, to be very similar for both the price and the child status variables. The sign patterns are identical, the patterns for coefficients that are significant are identical, and the numerical differences for the corresponding pairs of coefficient estimates are small in comparison with the differences between the *full:IV* and *restricted:IV* coefficient estimates.

Suppose there really are numerically important omitted variable biases in the coefficient estimates for the price and child status variables when the labor supply variables are left out of the budget-share equations. That is, suppose the labor supply variables do belong in the budget-share equations and, furthermore, that they are correlated with the included price and child status variables so that part of the effect of the labor supply variables will be spuriously captured by the price and child status coefficient estimates if the labor supply variables are omitted. If this is why the *full:IV* and *restricted:IV* estimates for the price and child status coefficients differ, then we would also expect the *full:direct* and *restricted:direct* price and child status coefficient estimates to differ for the same reason, though the pattern of the differences might be altered somewhat if the coefficient estimates for the *full:direct* and *restricted:direct* models are also affected by biases due to the endogeneity of the labor supply (and total expenditure) variables.

The above observations raise interesting questions. One might argue that the difference between the *full:direct* and *restricted:direct* estimates in Tables IV and V should be larger, given the omitted variable bias which must be present in the *restricted:direct* estimates. However, it could be countered that the biases due to endogeneity of some explanatory variables are sufficiently large as to dominate the omitted variable bias associated with excluding labor force variables. It is our contention that this is not likely. There are biases present due to excluding labor force variables, due to other omitted variables, and due to the endogeneity of explanatory variables in the *restriced:direct* estimation results. There is, however, no *a priori* reason to suspect that one of these biases is substantially larger than the other. Indeed, the focus of the BM paper, with its emphasis on the difference between the *full:IV* and *restricted:IV* results, and the arguments BM present indicate a pre-occupation with biases resulting from omitting the labor force variables. The orientation of their discussion and the results they present emphasise the importance of omitted variables bias and do not suggest that biases associated with endogeneity of explanatory variables might be particularly large. Furthermore, the results we present above regarding multicollinearity of the exogenous regressors with the generated instruments, and the correlation between the instruments and the variables they replace suggest that the IV results have larger MSE's than the non-IV results.

The foregoing discussion and the results in Tables IV and V suggests that the differences in the coefficient estimates for the *full:IV* and *restricted:IV* models reported by BM and also in the previous subsection of this paper are due primarily to a loss of precision due to the use of instrumental variables (particularly for the *full:IV* model) rather than to bias problems from the omission of the labor supply variables in the restricted model. This is supported by our observations that the generated regressor instruments are highly collinear with the other explanatory variables, and that the instruments themselves are generally weakly correlated with the variables they are used to instrument. Thus, our study replicates the BM rejection of the separability of commodity demands from labor supply, but raises questions about the conclusion that the omission of labor supply variables from budget-share equations will result in empirically important omitted variable bias problems. Rather, our results suggest that the inclusion of labor supply variables which are viewed as endogenous and hence are instrumented can lead to biases and losses of precision in estimation that are far more serious from the perspective of obtaining reliable point estimates of the coefficients of the budget-share equations than the omitted variable bias problems that might result from the omission of the labor supply variables. This speculative conclusion is further supported by the prediction results presented in the following section.

7 Prediction of Actual Budget-Shares

Muellbauer (1977) challenges researchers to examine the ability of their estimated demand systems to predict budget-shares for different types of households:

> Econometricians rarely report R^2's for budget-shares, often one suspects out of embarrassment ...
>
> *Muellbauer, 1977, p. 468.*

7.1 Our Results

Our demand share equations were estimated using pooled observations for households with different work status attributes (both worked: $d_m = 1$, $d_f = 1$; husband only worked: $d_m = 1$, $d_f = 0$; wife only worked: $d_m = 0$, $d_f = 1$; neither worked $d_m = 0$, $d_f = 0$). It is interesting to examine the capacity of our estimated equations for our different models to predict the pattern of actual average budget-shares for these different types of households for each commodity group.

Actual budget-shares are shown in row 1 for each commodity group panel of Table VI. Notice that the actual budget-shares for the $d_m = 1$, $d_f = 0$ and $d_m = 0$, $d_f = 1$ categories always lie between the $d_m = 1$, $d_f = 1$ and $d_m = 0$, $d_f = 0$ ones, except for the Utilities commodity group. That is, the actual expenditure shares for the families where only the man or only the woman worked lie between the expenditure shares for the two-worker families and the families where neither the man nor the woman worked.

Predicted budget-shares for the *full:direct*, *restricted:direct*, *restricted:IV* and *full:IV* models are shown in rows 2–5 of each commodity group panel. The values for the *full:direct* and *restricted:direct* models all follow essentially the same patterns as the actual values. This is not the case, however, for the *full:IV* model. For instance, for the four work status categories, the actual average budget-shares for Clothing are .164, .127,.128 and .096, with the largest budget-share for the $d_m = 1$, $d_f = 1$ category and the smallest for the $d_m = 0$, $d_f = 0$ category. The corresponding *full:direct* budget-shares are .164, .127, .133 and .092; and the *restricted:direct* shares are .162, .133, .134 and .095. These patterns are therefore *identical* for the actual budget-shares, compared to the pattern for the *full:direct* or *restricted:direct* predicted budget-shares. In contrast, for the *full:IV*, model the average budget-shares for Clothing are .169, .085, .355 and .199, with the largest budget-share for the $d_m = 0$, $d_f = 1$ category, and smallest for the $d_m = 1$, $d_f = 0$ category, and with a smaller budget share for the $d_m = 1$, $d_f = 1$ category than for the $d_m = 0$, $d_f = 0$ one. These predicted shares patterns are thus markedly different from the actual budget-shares.

The rankings of the budget-shares from largest (=1) to smallest (=4) over the four employment status categories are shown in parentheses in Table VI. The correspondences between the actual budget-shares and the predicted budget-shares for each empirical model can be summarized by the rank correlations, for which values are shown in the last column of Table VI. A rank correlation of .80 or higher indicates a significant positive relationship between the actual and predicted budget-share ranks for a one-tailed critical region of .05. Using this criterion, the predicted budget-shares for the *full:direct* and the *restricted:direct* models are significantly correlated with the actual budget-shares for all four commodity groups, while the only significant positive correlation for the *full:IV* model is for Services.

Based on the above discussion, it is clear that the *full:direct* and *restricted:direct* models yield more accurate budget-share predictions than the *full:IV* and *restricted:IV* models. It is true that, by construction, on average *full:direct* estimation will yield more precise predictions than *full:IV* estimation. However, the comparisons being made in Table VI relate to predictions for sub-groups of observations, which are sub-sets of the complete sample. In this setting, there is no *a priori* reason to expect the *full:direct* estimates to yield superior predictions. In particular, in these samples of predictions, there exist sub-sets of observations for which *full:IV* prediction is better than *full:direct* prediction. In the context of this study, however, we are interested in a comparison of predictive power in work status sub-group categories. For the type of classification of observations given the current sample, however, it turns out that *full:direct* predictions are superior to *full:IV* predictions on the basis of the rank correlation criterion.[4]

7.2 The BM Results

BM also show actual budget-shares and the predicted budget shares for their *full:IV* model. They show the actual and predicted figures in separate tables, and on different pages of their paper. Their discussion centers mostly on the predicted figures, and they do not compare their predicted budget-shares with their actual budget-shares. In Table VII, we show the BM actual and predicted figures together, making comparisons more convenient.

We see first of all that one of the generalizations that BM note for their predicted budget-shares does not hold for the actual budget-shares, examined by work status strata (both work, only man works, only woman works, neither works). For the predicted budget-shares, BM note:

> There are important differences in the budget share predictions for most goods and some of these are different from those for the means as reported in Table I [the actual budget-shares]. There is no strong pattern to the differences between the four strata; in particular we find no support for the idea that strata (1,0) and (0,1) [the two single worker strata] are convex combinations of strata (1,1) and (0,0) [the strata for families where both the man and woman work and for those where neither works].
>
> *Browning and Meghir, 1991, p. 940.*

That is, BM note that their predicted budget-shares for families where only the man or only the woman works do not lie between the shares for the two worker families and the families where

[4]When using microdata or panel data, in-sample predictive comparisons can often provide useful insights, in addition to the information obtained from out-of-sample prediction. See Heckman and Walker (1990, p. 1420), for a discussion of this issue.

neither the man nor the woman works. But the *actual* budget-shares for their British data *do* follow this pattern (except for the Other Goods category), just as both the actual budget-shares and the *full:direct* and *restricted:direct* predictions do for our Canadian data.

The rankings of the budget-shares from largest (=1) to smallest (=4) over the four employment status categories are shown in parentheses in Table VII. The corresponding rank correlations, given in the last column of the table, are significantly positive for only the Fuel and the Other Goods commodity categories. Thus, from a predictive perspective, BM's *full:IV* model performs very poorly, as is also the case for our *full:IV* model.

One reason for this could be a loss of precision in the estimation of the coefficients of the *full:IV* model for the same reasons we find reduced precision of our estimates in this study. That is, increased noise due to the use of poor instruments, and increased multicollinearity amongst the instrumented variables and the exogenous explanatory variables. If this is so, then their bias findings may also be highly imprecise.

8 Conclusions

The BM study is significant work, since it is one of the few pieces of research in applied demand analysis where the importance of labor supply effects on commodity demands is considered. That these effects are potentially substantial is not questioned in this paper. Indeed, our results support this finding of BM (and, more recently, Kaiser, 1993, using German data, and Nicol, 1995, using United States data). We do, however, believe that it makes a great deal of difference how labor force variables are treated when included in a demand model. The evidence we present indicates that, while such variables appear to be endogenous on the basis of standard test procedures, controlling for this endogeneity by using instrumental variables of the type BM use does not necessarily reduce bias to an acceptable level, and the concomitant loss of precision can contribute to interpretations which are quite misleading.

The foregoing observation seems to be particularly problematic in the context of labor force variables which are instrumented. In this paper, as in BM, we also instrument for total expenditure, and other variables which interact with total expenditure. This treatment does not seem to lead to such a large loss in precision as does instrumenting the labor force variables. This contention is borne out by analysing individual response coefficients, compared across estimation regimes, and also when comparing the respective predictive accuracy across those regimes as well. Indeed, one can see additional evidence of this in the study by Blundell, Pashardes and Weber (1993), who estimate a similar model with similar (FES) data to BM. In the Blundell, Pashardes and Weber (1993), no labor force variables are included, but total expenditure variables are instrumented. It is found in that case that prediction is fairly precise, relative to corresponding aggregate forecasts of micro share equation predictions.

To conclude, while it seems clear that labor supply variables are important determinants of demand, care must be exercised in the way in which these are handled when included in a demand system. It is not obvious, on the basis of our evidence, that one ought to instrument these variables when they are included, even if standard test results indicate their endogeneity. The purpose for which estimates (of parameters, predictions, or anything else) are to be used should be taken into account, to provide guidance on how to treat the labor force variables.

References

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

Browning, M. and C. Meghir, The Effects of Male and Female Labor Supply on Commodity Demands, *Econometrica* **59(4)**, 1991, 925–951.

Cragg, J.G. and S.G. Donald, Testing Identifiability and Specification in Instrumental Variable Models, *Econometric Theory* **9**, 1993, 222-240.

Deaton, A. and J. Muellbauer, An Almost Ideal Demand System, *American Economic Review* **70(3)**, 1980, 312-326.

Heckman, J.J. and J.R. Walker, The Relationship Between Wages and the Timing and Spacing of Births: Evidence from Swedish Longtitudinal Data, *Econometrica*, **58**, 1990, 1411-1441.

Kaiser, H., Testing Separability Between Commodity Demand and Labour Supply in West Germany, *Empirical Economics* **18**, 1993, 21–56.

Muellbauer, J., Testing the Barten Model of Household Composition Effects and the Cost of Children, *Economic Journal* **87**, 1977, 460-487.

Nakamura, A. and M. Nakamura, The Econometrics of Female Labor Supply and Children, *Econometric Reviews*, **11**, 1992, 1-71.

Nicol, C.J., Estimating a Third-Order Translog Using Canadian Cross-Sectional Micro-Data, *Canadian Journal of Economics* **22**, 1989, 543-560.

Nicol, C.J., Model Specification Issues in Consumer Demand Analysis Using United States Microdata. Department of Economics Working Paper No. 56, University of Regina, 1995.

Pollak, R.A., Conditional Demand Functions and Consumption Theory, *Quarterly Journal of Economics* **83**, 1969, 70-78.

Pollak, R.A., Conditional Demand Functions and Implications of Separability, *Southern Economic Journal* **37**, 1971, 423-433.

Rosenzweig, M.R. and T.P. Schultz, The Demand for and Supply of Births: Fertility and its Life Cycle Consequences, *American Economic Review*, **75**, 1985, 992-1015.

Rosenzweig, M.R. and K.I. Wolpin, Life-Cycle Labor Supply and Fertility: Causal Inferences from Household Models, *Journal of Political Economy*, **88**, 1980, 328-348.