Working Paper 95-57
Statistics and Econometrics Series 23
May 1996

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9849

HOUSEHOLD CHARACTERISTICS AND CONSUMPTION BEHAVIOUR: A NONPARAMETRIC APPROACH

Miguel A. Delgado and Daniel Miles.*

Abstract

In this paper we apply nonparametric methods in order to discuss some empirical aspects of household consumption behaviour. First, we study the differences in the consumption behaviour between household types. We find that, except for food, there are no clear significant differences. Secondly, we derive the functional form for the food Engel curve, using specification tests consistent in the direction of nonparametric alternatives. Finally, we use this specification to discuss the misleading conclusions that could be reached from a mechanic interpretation of the rejection of Hausman´s test, when applied to test the exogeneity of expenditure. The data is obtained from the Spanish Expenditure Survey 1980-81 and 1990-91.

Key Words
Consistent specification tests; Engel curves; Expenditure Endogeneity; Household characteristics; Nonparametric Estimation.

*Delgado, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid; Miles, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid. This article is based on research funded by Spanish "Dirección General de Investigación Científica y Técnica" (DGICYT), reference: PB92-0247.

Household Characteristics and Consumption Behaviour: A Nonparametric Approach

Miguel A. Delgado and Daniel Miles*

Universidad Carlos III de Madrid

May 14, 1996

Abstract

In this paper we apply nonparametric methods in order to discuss some empirical aspects of household consumption behaviour. First, we study the differences in the consumption behaviour between household types. We find that, except for food, there are no clear significant differences. Secondly, we derive the functional form for the food Engel curve, using specification tests consistent in the direction of nonparametric alternatives. Finally, we use this specification to discuss the misleading conclusions that could be reached from a mechanic interpretation of the rejection of Hausman's test, when applied to test the exogeneity of expenditure. The data is obtained from the Spanish Expenditure Survey 1980-81 and 1990-91.

Keywords: Engel curves; Househod characteristics; Nonparametric Estimation; Consistent Specification Tests; Expenditure Endogeneity.

JEL Classification Number: C14,C21,C52,D12.

* Article based on research funded by Spanish "Direccion General de Investigacion Cientifica y Tecnica" (DGICYT), reference number: PB92-0247.

1. INTRODUCTION

During the last years there has been a considerable amount of applied research on modeling household consumption behaviour. A motivation could be found on the need for interpreting demand parameters as inputs for evaluating the effectiveness of different economic policies, in particular, fiscal policies (Blundell, 1988; Baccouche and Laisney, 1989; Pollak and Wales, 1992; Nichele and Robin, 1995).

The estimated demand parameters are highly sensitive to the functional form of the Engel curve and to the way in which the relationship between consumption and household characteristics is specified, see, Blundell and Ray, (1984), Deaton, (1986) Blundell, (1988), Meghir and Robin, (1992), Pollak and Wales, (1992), among others.

Traditionally, a parametric functional form has been assumed for demand models. Misspecification of the functional form can produce inconsistent parameter estimators and, hence, it can yield invalid inferences and, in particular, misleading conclusions on the effects of household characteristics on consumption.

In this paper we apply different nonparametric inference techniques in order to investigate the relationship between household characteristics and consumption behaviour using the Spanish Expenditure Survey for the periods 1980-81 and 1990-91. The observations we consider are based on couples with the head of the household employed. We have constructed eleven subgroups depending on whether or not they have children, the age of the wife and the size of the city where they live. The goods we consider are food, alcohol, cloth, domestic fuel, recreation and other non-durable goods. Summary statistics on the sample used can be found in Table A1 and A2.

TABLE A1-A2 ABOUT HERE

In the next Section, we use a nonparametric approach to study how changes in characteristics affect consumption behaviour. We conclude that, except for food, consumption is not significantly affected by changes in household characteristics. Focusing on food consumption, in Section 3, we apply newly developed specification tests, consistent in the direction of nonparametric alternatives, in order to determine the correct functional form for the budget share Engel curve. The different tests are unable to reject the null hypothesis of a Working-Leser specification. In Section 4 we discuss possible reasons which can explain why in applied work total expenditure is found to be endogenous when a Hausman's Test is applied.

2. CHARACTERISTIC CHANGES AND SUBSTITUTION IN CONSUMPTION

In this section we study the relationship between household characteristics and consumption behaviour by estimating the substitution effects that take place in consumption when household characteristics change. Henceforth, these effects will be termed characteristic substitution effect (CSE).

We have a random sample of N observations $\left\{(W_{1i}, W_{2i}, ..., W_{Gi}, X_i, Z_i'), i = 1, ..., N\right\}$, where W_{mi} is the consumption budget share of good m, X_i denotes income (total expenditure) and Z_i is a vector of variables representing household characteristics. Suppose that Z_i only takes a number of discrete values producing H household groups. Then, the Engel curve is given by

$$m_m^h(x) = E[W_{mi}|X_i = x, Z_i = z^h] \quad m = 1, ..., G; \; h = 1, ..., H. \tag{1}$$

The CSE between two different households a and b is given by

$$s_m^{(a,b)}(x) = m_m^a(x) - m_m^b(x). \tag{2}$$

Notice that, by additivity,

$$\sum_{m=1}^{G} s_m^{(a,b)}(x) = 0 \quad a, b = 1, ..., H. \tag{3}$$

Therefore, when characteristics change, an increase in the consumption of one good must be compensated with a decrease in the consumption of some other good.

The Engel's curves in (2) are estimated using kernels, plugging-in an estimate of the optimal bandwidth minimizing the integrated mean square error (see, e.g. Härdle, 1990 and Wand and Jones, 1995). Then, $m_m^h(x)$ is estimated by

$$\widehat{m}_m^h(x) = \frac{\sum_{n=1}^{N} w_{mi} K((X_i - x)/h) I(Z_i = z^h)}{\sum_{n=1}^{N} K((X_i - x)/h) I(Z_i = z^h)},$$

where $I(A)$ is the indicator function of the event A and $K(u)$ is a kernel function (we have used the Gaussian density). The CSE is estimated by

$$\widehat{s}_m^{(a,b)}(x) = \widehat{m}_m^a(x) - \widehat{m}_m^b(x) \quad m = 1, ..., G; a, b = 1, ..., H.$$

The estimation results are presented in figures 1 to 4 for some particular subsamples and in Table A3 we summarize the range of variation of the CSE for those goods more sensitive to changes in household characteristics. Notice that the subsamples are differentiated only by one characteristic. In the figures, each line represents the difference of the estimated budget share of a given good for two household types.

In Figure 1 we observe that for families with the same expenditure level, no children and living in small towns, the older couples tend to increase the consumption of food and decrease the consumption of transport, recreation and the aggregate of other non-durables, as compared with the younger ones. This substitution process depends on the level of expenditure.

FIGURE 1 ABOUT HERE

Figure 1 also shows that goods such as alcohol, domestic fuel or cloth, play an insignificant role in the substitution process. Old and young couples behave similarly for those goods and this behaviour is not affected by the level of expenditure.

FIGURE 2-4 ABOUT HERE

Similar conclusions are derived from figures 2 to 4 and from the results summarized in Table A3. When we compare young couples, with or without children, the consumption of food of the former is higher than the latter comparatively, diminishing the aggregate of non-durables. The consumption of other goods do not seem to be affected by changes in household characteristics.

TABLE A3 ABOUT HERE

Our figures and Table A3 confirm Engel (1895) results when comparing households of different size for a wider class of household characteristics. Our data show that the food budget share is also affected, ceteris paribus, by changes in the size of the city where couples live or by changes in age. We also find that, for a given level of expenditure, consumption on goods such as cloth, alcohol or recreation is similar for all households.

Given these results on the food budget share, we report in figures 5 to 8 the food budget share Engel curves, estimated for different subsamples. These curves seem to be approximately linear. Among subsamples, Engel curves seem to differ only by a shift, which may justify the use of additive dummies in parametric settings (see e.g. Deaton et al., 1989, for Spanish cross-section data).

3. FUNCTIONAL FORM OF FOOD BUDGET SHARE

Food Engel curves estimated in the last Section are approximately linear. In this section we provide formal justification of the Working-Leser specification of the food Engel curve. Nested or non-nested tests employed in the empirical econometric literature are only consistent in the direction of certain alternatives which can be summarized by means of a given set of parameters. The test applied in this Section are consistent in the direction of general nonparametric alternatives.

The specification in the null hypothesis is given by

$$H_0 : m_m^h(x) = a_{m1}^*(z^h, \theta_1) + a_{m2}^*(z^h, \theta_2)\ln(x) + a_{m3}^*(z^h, \theta_3)(\ln(x))^2. \qquad (4)$$

where $a_{mj}^*(\cdot,\cdot)$ are given functions of household characteristics, z^h, and the vector of parameters, θ_j $j = 1, ..., m$.

Here we have implemented tests based on the comparison of a parametric fit and a nonparametric fit, based on kernel estimators (Horowitz and Härdle, 1994; Härdle and

Mammen, 1993; Ellison and Ellison, 1992). The performance of these tests depends on the choice of a bandwidth number for obtaining the kernel estimate. We report results for several bandwidth numbers choices. We have also implemented the test proposed by Stute (1995) which does to need to estimate the model in the alternative hypothesis and no bandwidth number has to be chosen. All test are presented in Appendix A.

The null hypothesis is generally not rejected for the household types considered. We present the results for the 1980 data, evaluated at 1990 prices in Table A4. For the 1990 sample, the null hypothesis is not rejected in the majority of subsamples. For the 1980 survey, only the Härdle and Mammen test rejects the null hypothesis for all selected bandwidth parameters in one case. These results are confirmed by Stute's test which does not reject the null in any of the subsamples.

TABLE A4-A5 ABOUT HERE

The quadratic term is not significantly different from zero, based on a t-test on the OLS coefficients, hence, the Working-Leser specification seem to be valid. In order to provide further evidence on the Working-Leser specification, we have also computed the average derivative estimator (ADE, Stoker ,1992; Härdle and Stoker, 1989; Powell, Stock and Stoker, 1989). The ADE estimates the parameter

$$\delta_m^h = E\left[\frac{\partial m_m^h(x)}{\partial x}\right] \quad h = 1, ..., H.$$

We employed the same bandwidth number that we used for the regression estimates. The results are reported in Table A6. The ADE and the OLS coefficients are quite similar, supporting the conclusions achieved with the specification tests on the validity of a Working-Leser specification.

TABLE A6 ABOUT HERE

Results in Table A6 also show that the coefficients between subsamples are quite similar for certain household groups, indicating that households characteristics seem to have only a shift translation effect on the Engel curves. Table A7 reports the results of an F-test on the OLS slope coefficients estimates in order to test that slope coefficients are identical. For some subsamples, the slope coefficients are significantly different, implying that some characteristics should interact with the expenditure variable in a parametric specification.

TABLE A7 ABOUT HERE

In the next Section, we use the above specification to discuss some misleading results that could appear when testing for exogeneity of total expenditure and adopting the mechanic instrumental variable approach when we the null is rejected.

4. EXOGENEITY OF EXPENDITURE AND SPECIFICATION TESTS

In the last section, we have concluded that the food Engel curves have the form

$$E(W_{mi} \mid X_i = x, Z_i = z_i^h) = \alpha_m^h + \beta_m^h \log(x), \quad h = 1, ..., H. \tag{5}$$

Therefore, total expenditure is, under this specification, exogenous. However, in many studies (see, for example, Deaton, (1986) and Blundell, (1988) among others), it has been argued that total expenditure is endogenous, and this argument has been supported by implementing a Hausman's test. Then, the result in the last section seems to be in contradiction with the generally argued endogeneity of total expenditure.

A possible reason for this contradiction could be found in the way characteristics variables enter the model, when running a regression for the whole sample, using dummy variables. Let the characteristics be represented by a dummy variable, z^i, one for each possible characteristic, $i = 1, .., C$, e.g., $z^1 = 1$ if household have children or 0 otherwise. The exhaustive model, taking into account all the possible interactive effects and the specification derived in the last section, is given by

$$W_{mi} = \alpha_m^r + X_i\beta_m^r + \sum_{c=1}^{C^*} Z_i^c(\alpha_m^c + X_i\beta_m^c) + \sum_{c=1,c<q}^{C^*} Z_i^c Z_i^q(\alpha_m^{cq} + X_i\beta_m^{cq}) +$$
$$\sum_{c=1,c<q<s}^{C^*} Z_i^c Z_i^q Z_i^s(\alpha_m^{cqs} + X_i\beta_m^{cqs}) ... + \prod_{c=1}^{C^*} Z_i^c(\alpha_m^{12...C} + X_i\beta_m^{12...C}) + u_i$$

where u_i is the error term, and r is the reference household, e.g., r could be couples with wife older than 35, living in small towns and without children, and C^* excludes those characteristics defining the reference household. The interaction terms of higher order, e.g. the interaction terms of second order are $z_n^c z_n^q$, which pick up the differential effects of a simultaneous change in characteristics, e.g. the simultaneous effect of a change from the reference household to a subsample of young couples with children will be captured by the second order interaction terms.

We have estimated the above exhaustive model for the food budget share and for the set of characteristics we were considering. We took as the reference household the old couples, without children and living in small towns. We detected that some interaction terms of second order between dummy variables, and between dummy variables and expenditure were significant. The second order interaction terms between characteristic dummy variables have a p-value of 0.023 and with total expenditure a p-value of 0.046 (Table A8). Based on an R-square selection criteria, we have selected the model presented in Table A9. As could be expected by the figures of Section 2, the additive dummies have a relative important impact on the food budget share, while the interaction terms between dummies and expenditure, although significant, are of relatively less relevance. Notice also that the interaction term between age and number of children is significant, capturing the simultaneous change from the reference household to young couples with children. In empirical specifications, it is not common to include the interaction terms which capture the effect of simultaneous changes in characteristics, which could introduce an omitted variable problem. The omission of these interaction terms could induce to reject the null hypothesis when using the Hausman test.

TABLE A8 A9 ABOUT HERE

Another reason for rejecting the null hypothesis can be found in the validity of the instrumental variables used. We had applied the Hausman test to the specification (5) for each subsample. Although the null should not be rejected, given the results of the previous section, we found that the null is sometimes rejected, depending on the set of instrumental variables used (Table A10). Then, the same set of instrumental variables do not need to be valid for different subsamples, and therefore, this cast doubts on the use of the same set of instrumental variables for the whole sample, as it is common in the empirical literature. Finally, another argument for the contradiction between the results of the specification tests and Hausman test could be found in the overrejection problem of the Hausman's test.

TABLE A10 ABOUT HERE

To conclude, the rejection of the null when applying a Hausman's test could arise from reasons other than the total expenditure endogeneity. This implies that adopting the mechanic instrumental variable approach could seriously bias the estimates.

5. CONCLUSIONS

This paper has applied nonparametric regression techniques for studying the relationship between consumption and household characteristics based on the Spanish Expenditure Survey. We have seen that food consumption is affected by household characteristics but other goods such as cloth, recreation and alcohol are not specially affected. Specification tests, consistent in the direction of nonparametric alternatives, support the Working-Leser model for the food Engel curve for the different household groups considered. Finally, we have shown that the null hypothesis, when applying Hausman's Test, can be rejected due to other reasons, like the way household characteristics are introduced, the validity of the instrumental variables, or the bad performance of Hausman's Test in small samples, rather than to total expenditure endogeneity.

REFERENCES

[1] Baccouche, R. and Laisney, F (1989) "Simulation of VAT reforms for France using cross-section data" in Florens, J.P., Ivaldi, M., Laffont, J.J. and Laisney, F. (eds) *Microeconometric: Surveys and Applications.* Basil Blackwell. Oxford.

[2] Banks, J. Blundell, R. and Lewbel, A. (1994) "Quadratic Engel Curves, Indirect Tax Reform and Welfare Measurement" University College London, Discussion Paper.

[3] Bierens, H.J. and Pott-Buter, H.A. (1990) "Specification of household Engel Curves by Nonparametric Regression", *Econometric Reviews*, 9(2), 123-184.

[4] Blundell, R. (1990) "Specification of Household Expenditure Functions and Equivalence Scales by Nonparametric Regression" Comment, *Econometric Reviews*, 9(2), 185-187.

[5] Blundell, R. (1988) "Consumer Behaviour: Theory and Empirical Evidence-A Survey", *The Economic Journal*, 98, March, 16-65.

[6] Blundell, R. and Ray,R. (1984) "Testing for Linear Engel Curves and Additively Separable Preferences using new Flexible Demand Systems", *Economic Journal*, 94, 800-811.

[7] Deaton, A. (1986) "Demand Analysis" in *Handbook of Econometrics*, ed. by Z. Griliches and M.D. Intrilligator, Elsevier Science Publishers.

[8] Deaton, A. and Muellbauer, J. (1980) *Economics and Consumer Behaviour.* Cambridge: Cambridge University Press.

[9] Deaton, A., Ruiz-Castillo, J. and Thomas, D. (1989) "The Influence of Household Composition on Household Expenditure Patterns: Theory and Spanish Evidence", *Journal of Political Economy,* 97,1, 179-200.

[10] Ellison G. and Ellison, S.F. (1992) "A Nonparametric Residual-based Specification Test: Asymptotic, Finite Sample, and Computational Properties" manuscript.

[11] Fry, V. and Pashardes, P. (1992) "An Almost Ideal Quadratic Logarithmic Demand System for the Analysis of Micro Data" Institute for Fiscal Studies

[12] Gorman, W.M. (1981), "Some Engel Curves", in *The Theory and Measurement of Consumer Behaviour*, ed. by A. Deaton, Cambridge: Cambridge University Press.

[13] Härdle, W. (1990) *Applied Nonparametric Regression.* Cambridge: Cambridge University Press.

[14] Härdle, W. and Mammen, E. (1993) "Comparing Nonparametric Versus Parametric Regression Fits", *The Annals of Statistics*, 21(4), 1926-47.

[15] Härdle, W. and Stoker, T.M. (1989) "Investigating Smooth Multiple Regression by the Method of Average Derivatives" *Journal of the American Statistical Association*, 84, 986-995.

[16] Hausman, J.A. (1986) " Specification and Estimation of Simultaneous Equation Models", Griliches, Z. and Intriligator M. (eds), *Hanbook of Econometrics.* North-Holland.

[17] Hausman, J.A., Newey, W. and Powell, J.L. (1995) "Nonlinear Errors in Variables Estimation of Some Engel Curves" *Journal of Econometrics*,65, 205-33.

[18] Horowitz, J.L. and Härdle, W. (1994) "Testing a Parametric Model Against a Semiparametric Alternative", Econometric Theory, 10, 821-48.

[19] Lewbel, A. (1991) "The Rank of Demand Systems: Theory and Nonparametric Estimation" *Econometrica,* 59, 711-30.

[20] Meghir, C. and Robin, J.M. (1992), "Frequency of Purchase and Estimation of Demand Systems" *Journal of Econometrics*, 53, 53-85.

[21] Nicol, C.J. (1993) "An Empirical Comparison of Nonparametric and Parametric Engel Functions", *Empirical Economics,*18:233-249.

[22] Nichele, V. and Robin, J.M. (1995) "Simulation of Indirect Tax Reforms Using Pooled Micro and Macro French Data" *Journal of Public Economics,* 56, 225-244.

[23] Pollak R. and Wales, T. (1992) *Demand System Specification and Estimation.* Oxford: Oxford University Press.

[24] Spencer, D.E. and Berk, K. (1981) "A Limited Information Specification Tets", *Econometrica*, 49,4, 1079-1085.

[25] Stoker, T.M. (1991) *Lectures on Semiparametric Econometrics.* CORE Lecture Series. CORE Foundation, Universite Catholique de Louvain.

[26] Stute, W. (1995) " Nonparametric Model Checks for Regression" forthcoming in *The Annals of Statistics.*

[27] Stute, W., Manteiga, W. and Presedo, M. (1995) "Bootstrap Approximations in Model Checks for Regression", forthcoming in Journal of the American Statistical Association.

[28] Wand, M.P. and Jones, M.C. (1995) "*Kernel Smoothing*" Chapman & Hall.

APPENDIX A

Let (X, Y) denote a $R^d \times R$ random vector and assume that Y is integrable so that the regression function

$$m(x) = E[Y \mid X = x] \quad x \in R^d,$$

is well defined. Parametric modeling assumes that m belongs to a given family

$$\mathcal{M} = \{m(\cdot, \theta) : \theta \in \Theta\},$$

where $\Theta \subset R^p$, assuming that $m(x) = m(x, \theta_o)$ for some true parameter θ_o. Statistical inference based on $\mathcal{M}$ should be accompanied by a test for

$$H_0 : m \in \mathcal{M} \textit{ versus } H_1 : m \notin \mathcal{M}.$$

Statistics for testing this hypothesis:

Stute (1995): Let

$$R_n^1(x) = n^{-1/2} \sum_{i=1}^{n} I(X_i \leq x) \left[Y_i - m(X_i, \widehat{\theta})\right]$$

where $I(A)$ is the indicator function of event A, $\widehat{\theta}$ is a $\sqrt{n}$ consistent estimator of θ. The statistic is given by

$$W_n^2 = \sigma_n^{-2} \int \left[R_n^1(x)\right]^2 F_n(dx)$$

where σ_n^2 can be estimated by any consistent estimate of σ^2. To compute the critical values a wild bootstrap procedure proposed by Stute et al. (1995) is used.

Härdle and Mammen (1993):

$$T_n = nh^{d/2} \int (\widehat{m}_h(x) - \mathcal{K}_{h,n} m(x, \widehat{\theta}))^2 \pi(x) dx$$

where h is the bandwidth parameter, $\widehat{m}_h(x)$ is the nonparametric regression, $\mathcal{K}_{h,n} m(x, \widehat{\theta})$ is the nonparametric regression of the parametric fit and $\pi(x)$ is a weighting function, equal to zero at the 5% observations at the tails and one otherwise. The critical value is approximated by wild bootstrap.

Ellison and Ellison (1992):

$$T_n = \frac{\left(Y_i - m(x,\widehat{\theta})\right)' W_n \left(Y_i - m(x,\widehat{\theta})\right)}{\sqrt{2}\widehat{\sigma}^2 s(W_n)}$$

where W_n is a weighting matrix with elements

$$w_{ij} = K((x_i - x_j)/h) / \sum_{t \neq i} K((x_i - x_j)/h) \quad if\ i \neq j,$$

where $K(u)$ is a gaussian kernel,

$$s(W_n) = \left(\sum_{i,j} w_{ij}^2\right)^{1/2},$$

and $\widehat{\sigma}^2 = n^{-1}\sum_{i=1}^{n}\left[Y_i - m(X_i,\widehat{\theta})\right]^2$. They show that $T_n \rightarrow \mathcal{N}(0,1)$, with a finite sample correction bias.

Horowitz and Härdle (1994):

The model to be tested in this case is a semiparametric model, where the hypothesis are given by

$$\begin{aligned} H_0 &: E\left[Y \mid v(x,\theta) = v\right] = F(v)\ versus \\ H_1 &: E\left[Y \mid v(x,\theta) = v\right] = H(v) \end{aligned}$$

where $v(\cdot)$ and $F(\cdot)$ are assumed to be known and H is unknown. Then, given

$$T_n = h^{1/2}\sum_{i=1}^{n} \pi(v(x,\widehat{\theta}))\left\{\left(Y_i - F\left(v(x,\widehat{\theta})\right)\right) \times \left(\widehat{F}_n\left(v(x,\widehat{\theta})\right) - F\left(v(x,\widehat{\theta})\right)\right)\right\}$$

where $\widehat{F}_n\left(v(x,\widehat{\theta})\right)$ is the nonparametric regression of $F\left(v(x,\widehat{\theta})\right)$ and $\pi(v(x,\widehat{\theta}))$ is a weighting function, that as before, excludes the 5% observations at the tails. The null can be accepted or rejected at ζ level according to whether $(T_n/\widehat{\sigma})$ exceeds the $1-\zeta$ quantile of the standard normal distribution, being a one sided test. For the formula for estimating the variance, see Horowitz and Härdle, 1994.

TABLE A1: Spanish Family Expenditure Survey, 1980-81;1990-91. Sample: Head of Household Employed and between 20 and 65 years old. Children: number of children, 0,1-2, and 3 or more. Age of Partner, younger or older 35 years old. Size of the City, larger or smaller 50th. habitants.

				Observations	
Subsample	Children	Age Partner	Size City	1980	1990
1	0	more 35	less 50 th.	514	379
2	0	more 35	more 50 th.	508	197
3-4	0	less 35	—	407	378
5	1-2	more 35	less 50 th.	765	862
6	1-2	more 35	more 50 th.	1047	625
7	1-2	less 35	less 50 th.	942	1275
8	1-2	less 35	more 50 th.	1548	778
9	3 or +	more 35	less 50 th.	613	381
10	3 or +	more 35	more 50 th.	802	222
11	3 or +	less 35	less 50 th.	353	230
12	3 or +	less 35	more 50 th.	501	109

TABLE A2: Summary Statistic of the Budget Share and Log of Total Nondurable Expenditure for 1980-81 and 1990-91 samples.

1980				
Goods	Mean	St.Dev	Min.	Max.
Food	0.329	0.142	0.004	0.920
Alcohol	0.013	0.020	0	0.251
Cloth	0.080	0.063	0	0.645
Domestic Fuel	0.031	0.024	0	0.367
Transport	0.103	0.089	0	0.454
Recreation	0.039	0.049	0	0.685
Other Non-Durables	0.403	0.120	0.04	0.948
Total Log Expenditure	13.73	0.601	11.19	16.31
1990				
Food	0.254	0.118	0.003	0.781
Alcohol	0.009	0.017	0	0.267
Cloth	0.098	0.088	0	0.774
Domestic Fuel	0.036	0.028	0	0.454
Transport	0.124	0.119	0	0.829
Recreation	0.041	0.046	0	0.470
Other Non-Durables	0.438	0.140	0.0325	0.956
Total Log Expenditure	14.64	0.508	12.43	16.48

TABLE A3: Summary of those goods which are more representative in the Characteristic Substitution Effect estimation. F: food; T:transport; C:cloth; R:recreation O: Other non-durables; NC: not clear, when no good dominates in the substitution process. Min and Max are the minimum and maximum of the substitution range.

COMPARISON	1980	MIN	MAX	1990	MIN	MAX
Old Couples,No Child, Small vs Big Cities	F-O	-0.06	0.06	NC	-0.06	0.02
No Child, Small Town, Old vs Young Couples	F-O	-0.18	0.14	F-O	-0.11	0.11
Old Couples, Small Town, With vs Without Child	F-O-C	-0.05	0.04	F-O	-0.04	0.04
No Child, Big Cities, Old vs Young Couples	F-T-C-O	-0.09	0.07	F-T-C-O	-0.10	0.13
Old Couples, Big Cities, With vs Without Child	F-O	-0.05	0.04	NC	-0.05	0.08
Young Couples,No Child, Small vs Big Cities	F-O	-0.09	0.07	NC	-0.06	0.09
Young Couples, Small Town, With vs Without Child	F-O	-0.13	0.15	F-O	-0.11	0.13
Young Couples, Big Cities, With vs Without Child	F-O	-0.08	0.1	F-T-O	-0.14	0.10
Old Couples,With Child, Small vs Big Cities	F-O	-0.06	0.04	F-C-O	-0.05	0.03
With Child, Small Town, Old vs Young Couples	F-T	-0.06	0.06	F-T-O	-0.03	0.05
With Child, Big Cities, Old vs Young Couples	F-O	-0.04	0.04	F-T-O	-0.05	0.03
Young Couples,With Child, Small vs Big Cities	F-O	-0.05	0.03	O-C	-0.04	0.03

TABLE A4: Results of Ellison and Ellison, Horowitz and Härdle, and Härdle and Mammen Specification Test for 1980-81 data. The Ellison and Ellison statistics is asymptotically distributed as a Standard Normal. The Horowitz and Härdle statistics is asymptotically distributed as a Standard Normal and the test is a one side test. Härdle and Mammen P-values were obtained by Wild Bootstrap with 200 resamples. Stute P-values were obtained by Wild Bootstrap with 500 resamples.

	Ellison and Ellison		Horowitz and Härdle		Härdle and Mammen				Stute	
Sample	h=0.15	h=0.2	h=0.15	h=0.2	h=0.15	P-v	h=0.2	P-v	S	P-v
1	-0.741	-0.288	-3.51	-3.794	0.069	0.87	0.056	0.80	0.516	0.46
2	-0.015	-0.059	-2.34	-2.98	0.079	0.43	0.057	0.43	0.573	0.40
3-4	-0.74	-0.11	-4.29	-6.51	0.022	0.76	0.017	0.074	0.616	0.37
5	-0.989	0.06	-1.85	-1.155	0.063	0.48	0.042	0.63	0.556	0.45
6	0.154	2.09	-0.98	-4.05	0.098	0.02	0.088	0.03	0.583	0.50
7	1.027	0.434	-2.10	-3.65	0.057	0.36	0.041	0.41	0.592	0.43
8	0.324	0.452	-8.33	-13.22	0.027	0.66	0.022	0.58	0.600	0.42
9	2.272	-0.20	-1.43	-0.57	0.048	0.72	0.038	0.66	0.560	0.45
10	-0.423	-0.560	-3.06	-3.76	0.021	0.84	0.017	0.78	0.568	0.40
11	1.327	-0.17	-0.69	0.24	0.054	0.58	0.041	0.54	0.553	0.39
12	0.37	-0.22	-2.24	-3.94	0.016	0.9	0.016	0.86	0.617	0.34

TABLE A5: Testing the Significance of the Expenditure Quadratic Term in the Food Engel Curve.

SAMPLE	P-VALUE 1980	P-VALUE 1990
1	0.2067	0.9719
2	0.3487	0.2254
3-4	0.8032	0.2024
5	0.7930	0.8603
6	0.0407	0.4574
7	0.7225	0.3402
8	0.3053	0.7551
9	0.3777	0.7642
10	0.7375	0.9760
11	0.0540	0.5597
12	0.1130	0.6466

TABLE A6: Average Derivative Estimator and OLS estimation of the slope of the Food Engel Curve Working-Leser Specification.

	1980(1990 prices)		1990	
Subsample	OLS	ADE	OLS	ADE
1	-0.169 (0.010)	-0.172 (0.014)	-0.114 (0.011)	-0.115 (0.014)
2	-0.156 (0.011)	-0.157 (0.016)	-0.137 (0.014)	-0.117 (0.0212)
3-4	-0.124 (0.010)	-0.126 (0.015)	-0.094 (0.001)	-0.113 (0.012)
5	-0.175 (0.010)	-0.184 (0.013)	-0.146 (0.006)	-0.145 (0.010)
6	-0.167 (0.010)	-0.152 (0.011)	-0.124 (0.010)	-0.118 (0.012)
7	-0.170 (0.010)	-0.160 (0.013)	-0.134 (0.005)	-0.143 (0.008)
8	-0.157 (0.005)	-0.153 (0.009)	-0.118 (0.007)	-0.123 (0.010)
9	-0.194 (0.008)	-0.194 (0.014)	-0.162 (0.010)	-0.156 (0.016)
10	-0.178 (0.007)	-0.186 (0.013)	-0.142 (0.012)	-0.130 (0.016)
11	-0.186 (0.013)	-0.177 (0.022)	-0.158 (0.013)	-0.132 (0.022)
12	-0.186 (0.009)	-0.196 (0.016)	-0.135 (0.0164)	-0.124 (0.0232)

TABLE A7: Rejection results of testing the significance difference between the parameters of Table A6. Results are given in terms of the subsample for which the difference is significant

	1980 Size		1990 Size	
Subsample	5%	10%	5%	10%
1	4	4,9	5,9,11	5,7,9,11
2		4,9,10,11,12	4	4
3-4	ALL,except 2	ALL	ALL, except 1	ALL, except 1
5	4	4,8	1,4,6,8	1,4,6,8
6	4,9	4,9	4,5,9,11	4,5,9,11
7	4,9	4,9	4,9	1,4,8,9,11
8	4,9,10,11,12	4,5,9,10,11,12	4,5,9,11	4,5,7,9,10,11
9	2,4,6,7,8	1,2,4,6,7,8	1,4,6,7,8	1,4,6,7,8
10	4,8	2,4,8	4	4,8
11	4,8	2,4,8	1,4,6,8	1,4,6,7,8
12	4,8	2,4,8	4	4

TABLE A8: Estimation of the food Engel Curve for the whole sample including higher order interaction terms between dummies and between dummies and expenditure, based on 1980-81 data.

Variables	Coefficient	T-Value
C	2.802	18.08
AGE	-0.889	-3.61
SIZEC	-0.460	-2.25
CHILDREN	0.003	0.057
AGESIZE	0.508	1.635
AGECHIL	0.306	2.878
SIZECHI	0.098	1.188
ASC	-0.178	-1.293
EXPEND	-0.182	-15.81
AGEEXP	0.060	3.295
SIZEXP	0.032	2.161
CHIEXP	0.001	0.2443
ASEXP	-0.038	-1.672
ACHEXP	-0.021	-2.700
SICHEXP	-0.007	-1.175
ASCEXP	0.013	1.322

Testing Joint Significance of interaction terms:

Dummy Variables (AGESIZE,AGECHIL,SIZEHI,ASC): **p-value 0.0236**

Dummy Variables and Expenditure (ASEXP, ACHEXP, SICHEXP, ASCEXP): **p-value 0.0461**

VARIABLES USED IN TABLE A8

AGE: Dummy: 1 partner younger 35; 0 otherwise.

SIZEC: Dummy: 1 if city is bigger 50.000 habitants; 0 otherwise.

CHILDREN: Number of Children.

AGESIZE: Product of AGE and SIZEC.

AGECHIL: Product of AGE and CHIL

SIZECHI: Product of SIZE and CHILDREN

ASC: Product of AGE, SIZE and CHILDREN

EXPEND: Total Non-durable Expenditure

AGEEXP: Product AGE and EXPEND

SIZEXP: Product SIZEC and EXPEND

CHIEXP: Product CHILDREN and EXPEND

ASEXP: Product AGE SIZE EXPEND

ACHEXP: Product AGE CHILDREN EXPEND

SICHEXP: Product SIZEC CHILDREN EXPEND

ASCEXP: Product AGE SIZEC CHILDREN EXPEND

TABLE A9: Model derived from that of Table A8 based on an R-square selection criteria

Variables	Coefficient	T-Value
C	2.555	25.33
AGE	-0.246	-2.95
SIZEC	-0.205	-2.39
CHILDREN	0.108	3.374
AGECHIL	0.026	10.31
EXPEND	-0.164	-22.01
AGEEXP	0.012	1.960
SIZEXP	0.014	2.206
CHIEXP	-0.006	-2.79

TABLE A10: Results of applying Hausman Test for different subsets of Instrumental Variables and for each of the Subsamples with 1980-81 data. R: reject the null. N: non-reject.

INSTRUMENTAL VARIABLE	1	2	3-4	5	6	7	8	9	10	11	12
Subset 1	R	R	R	N	N	R	R	N	N	N	N
Subset 2	R	N	N	N	N	R	R	N	R	R	N
Subset 3	R	N	N	N	N	R	R	N	R	R	N
Subset 4	R	R	R	R	R	R	R	R	R	R	R
Subset 5	R	R	R	R	R	R	R	R	R	N	N

Note: The subset of Instrumental Variables where define as follow:

Subset 1: husband age and husband age square.

Subset 2: Subset 1 + Semester dummy variable

Subset 3: Subset 2 + Partner Age divided and Multiply by Husband Age

Subset 4: Subset 3 + Education + Owning a Car + Blue Collar

Subset 5: Subset 4 + Log Income.

FIGURE 1

CSE Estimates with 1980 Data of
Couples with Partner Older than 35 vs Younger than 35
with no children and living in a Small Town

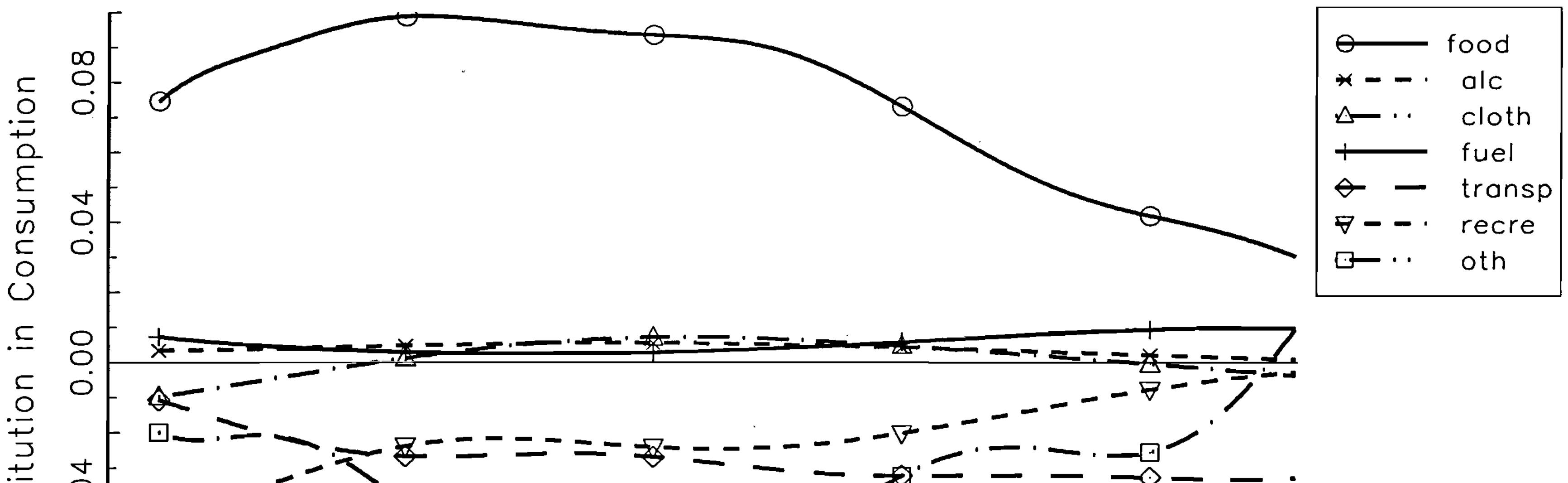


FIGURE 2

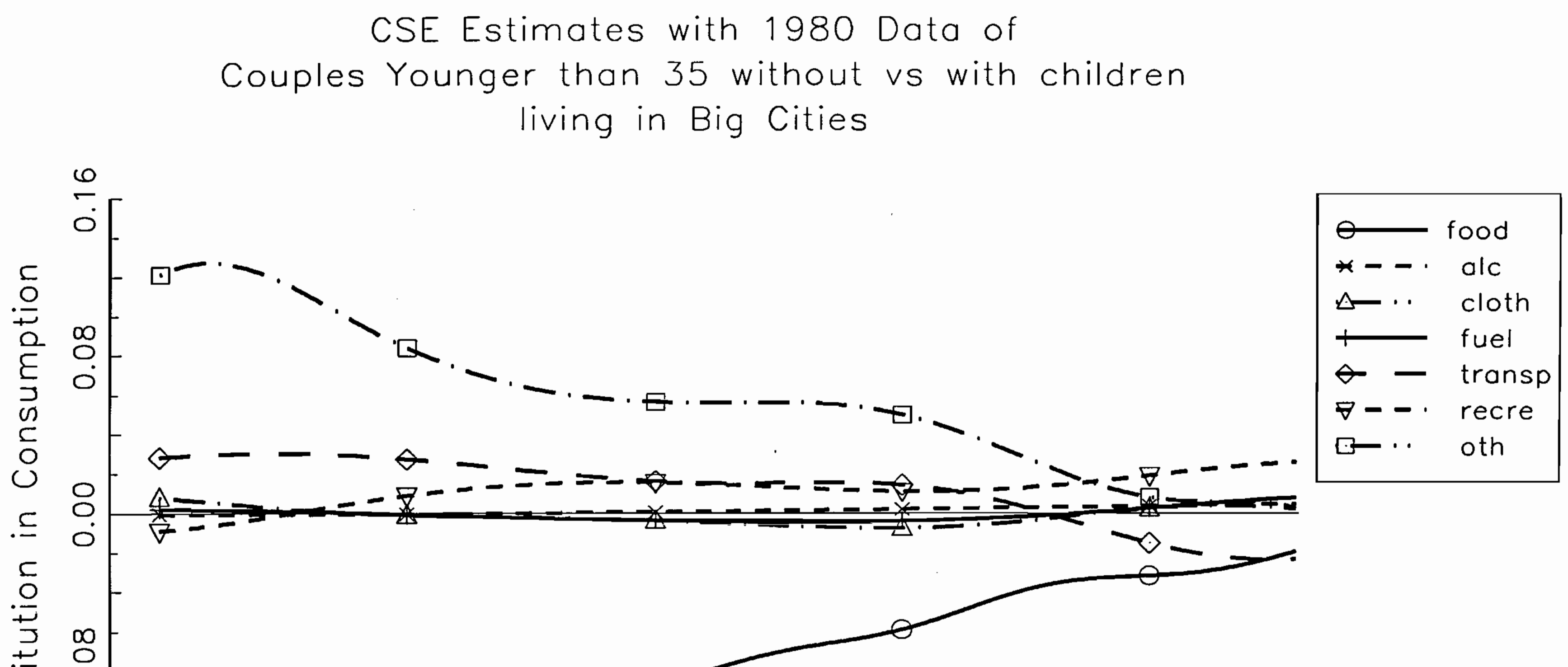

CSE Estimates with 1980 Data of
Couples Younger than 35 without vs with children
living in Big Cities
0.16
0.08
0.00
08
itution in Consumption
food
alc
cloth
fuel
transp
recre
oth

FIGURE 3

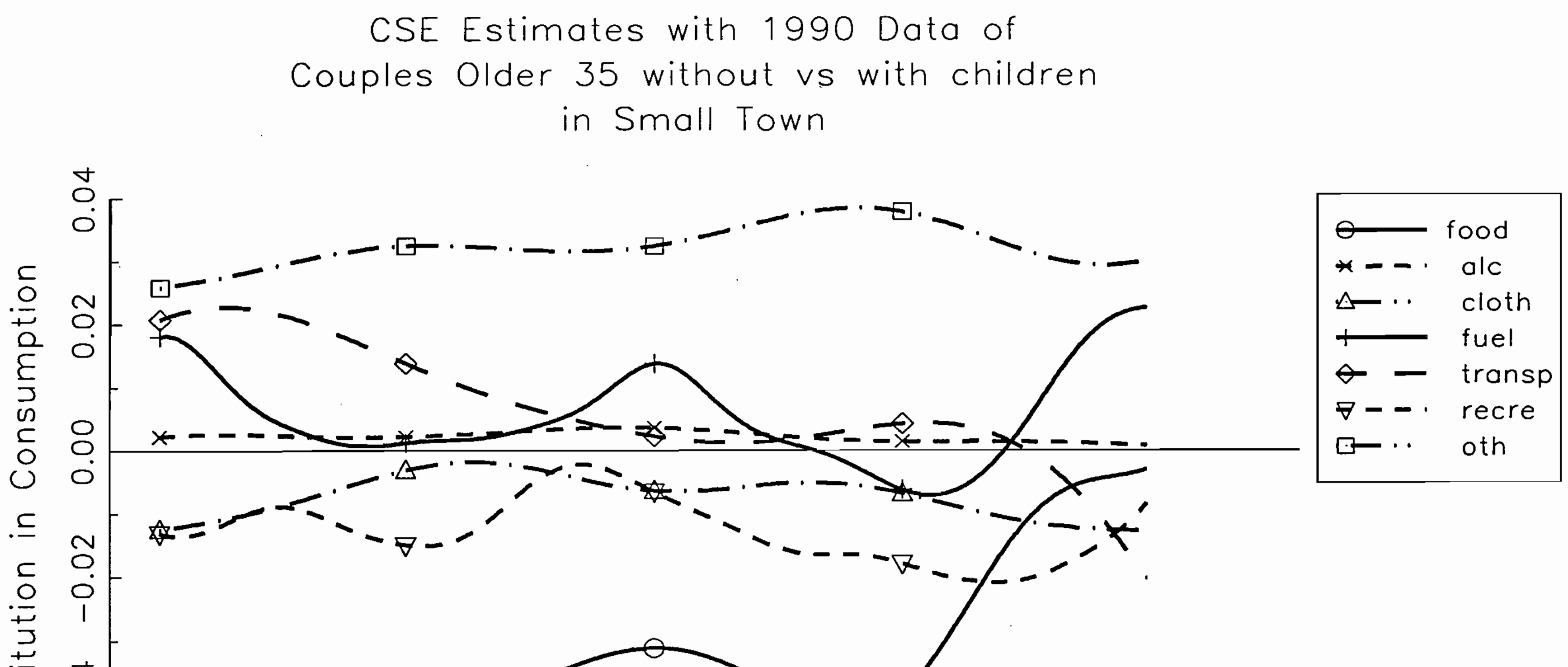

CSE Estimates with 1990 Data of
Couples Older 35 without vs with children
in Small Town
0.04
0.02
0.00
-0.02
itution in Consumption
food
alc
cloth
fuel
transp
recre
oth

FIGURE 4

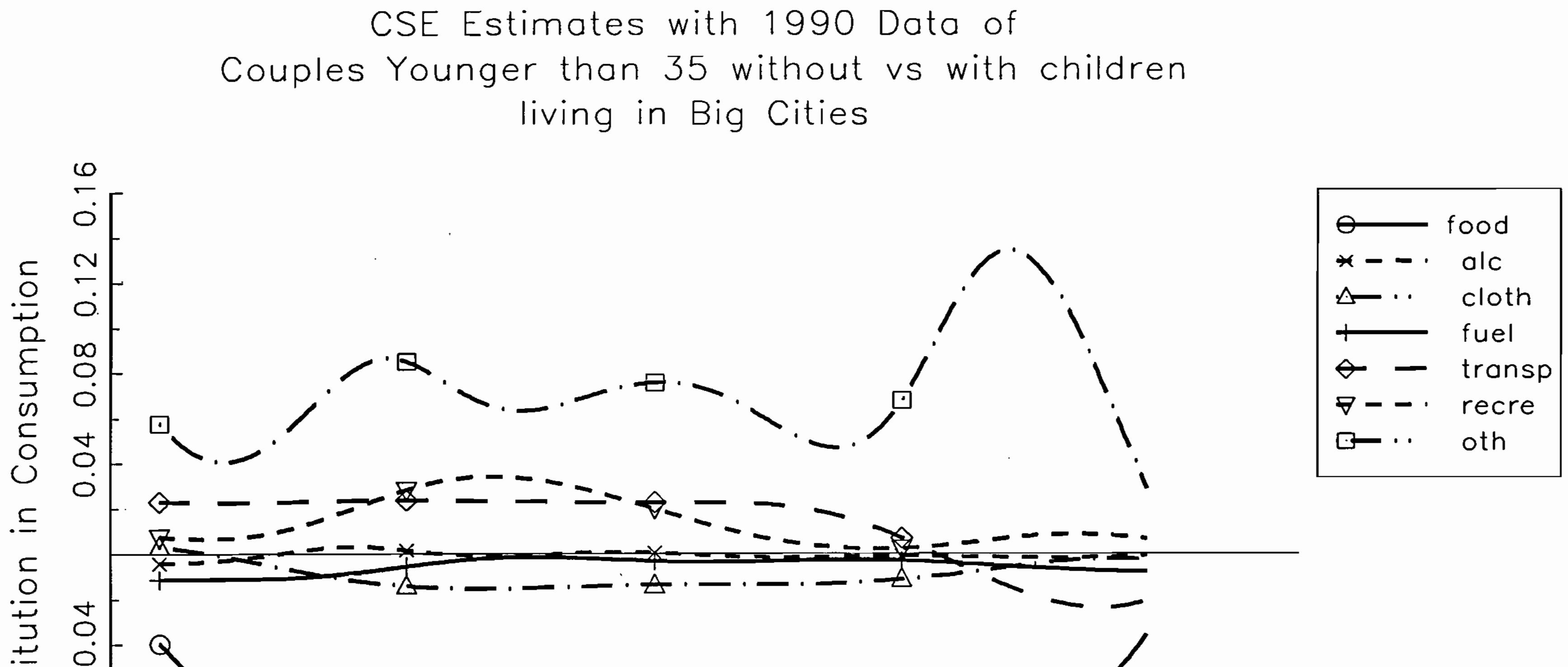

CSE Estimates with 1990 Data of
Couples Younger than 35 without vs with children
living in Big Cities
itution in Consumption
0.16
0.12
0.08
0.04
0.04
food
alc
cloth
fuel
transp
recre
oth

FIGURE 5

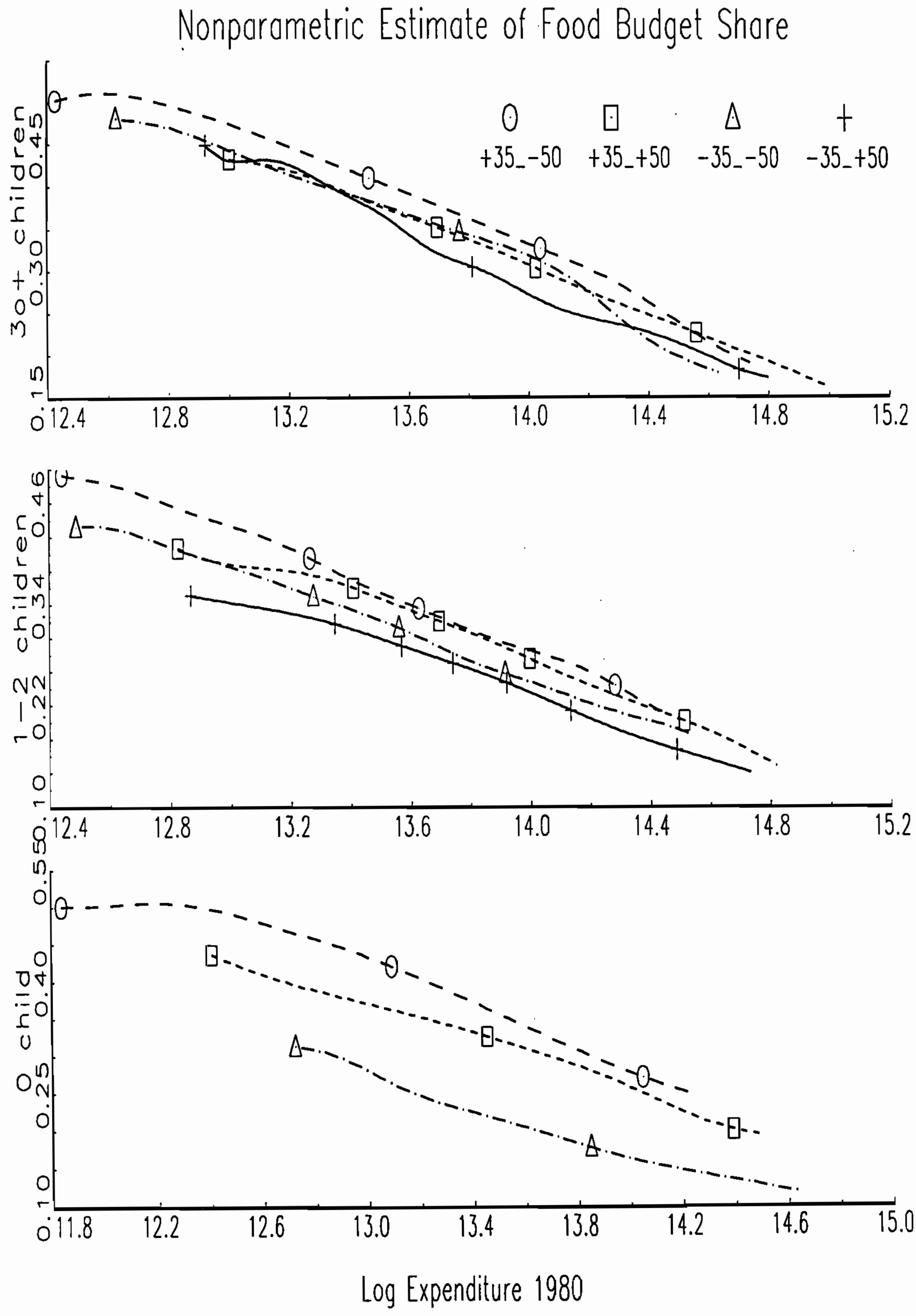
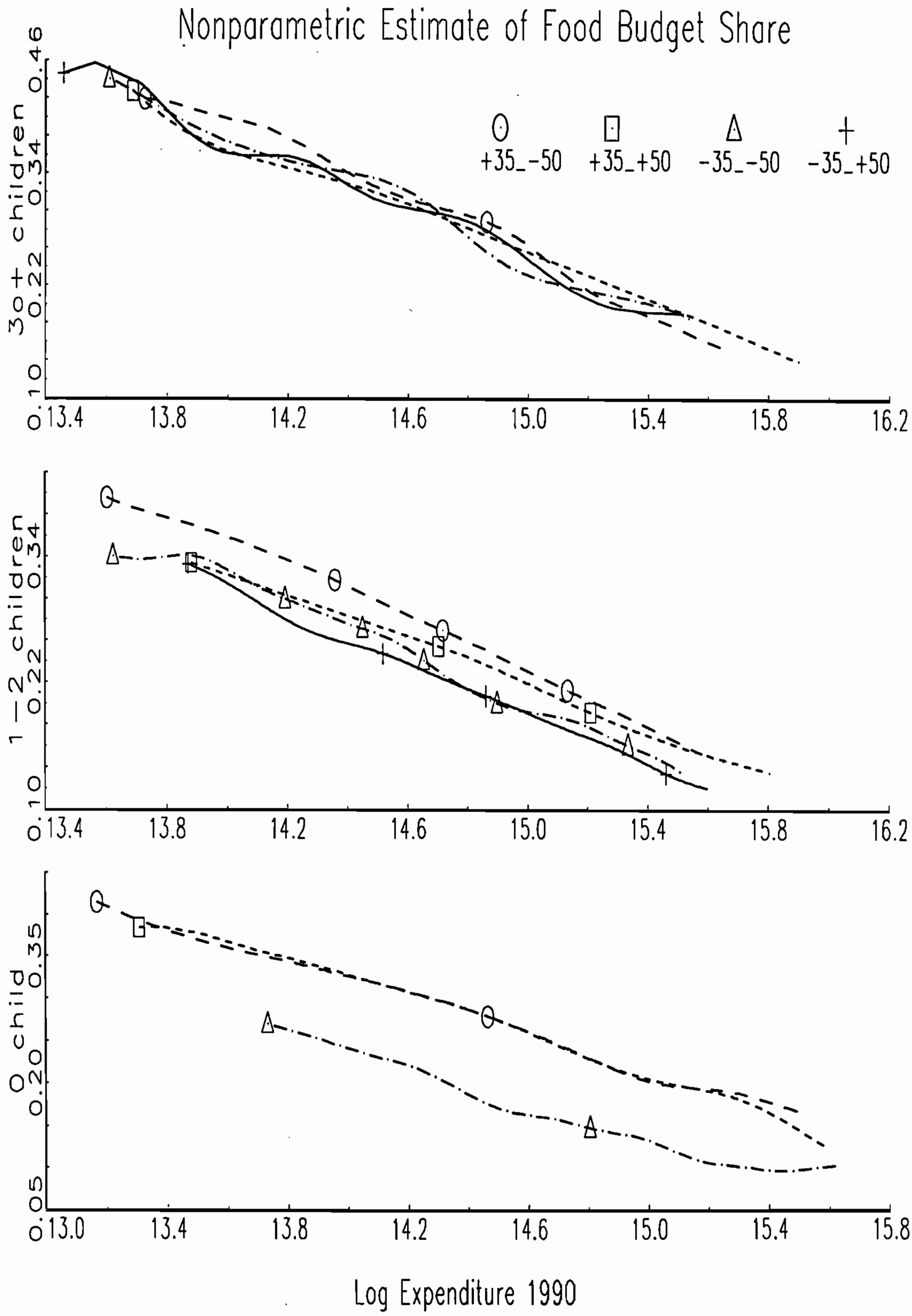

Nonparametric Estimate of Food Budget Share
+35_-50
+35_+50
-35_-50
-35_+50
3o+ children
0.15
0.30
0.45
12.4
12.8
13.2
13.6
14.0
14.4
14.8
15.2
1-2 children
0.10
0.22
0.34
0.46
0 child
0.10
0.25
0.40
0.55
11.8
12.2
12.6
13.0
13.4
13.8
14.2
14.6
15.0
Log Expenditure 1980

FIGURE 6



Nonparametric Estimate of Food Budget Share
+35_-50
+35_+50
-35_-50
-35_+50
3o+ children
0.46
0.34
0.22
0.10
13.4
13.8
14.2
14.6
15.0
15.4
15.8
16.2
1-2 children
0.34
0.22
0.10
13.4
13.8
14.2
14.6
15.0
15.4
15.8
16.2
0 child
0.35
0.20
0.05
13.0
13.4
13.8
14.2
14.6
15.0
15.4
15.8
Log Expenditure 1990

FIGURE 7

Nonparametric Estimate of Food Budget Share

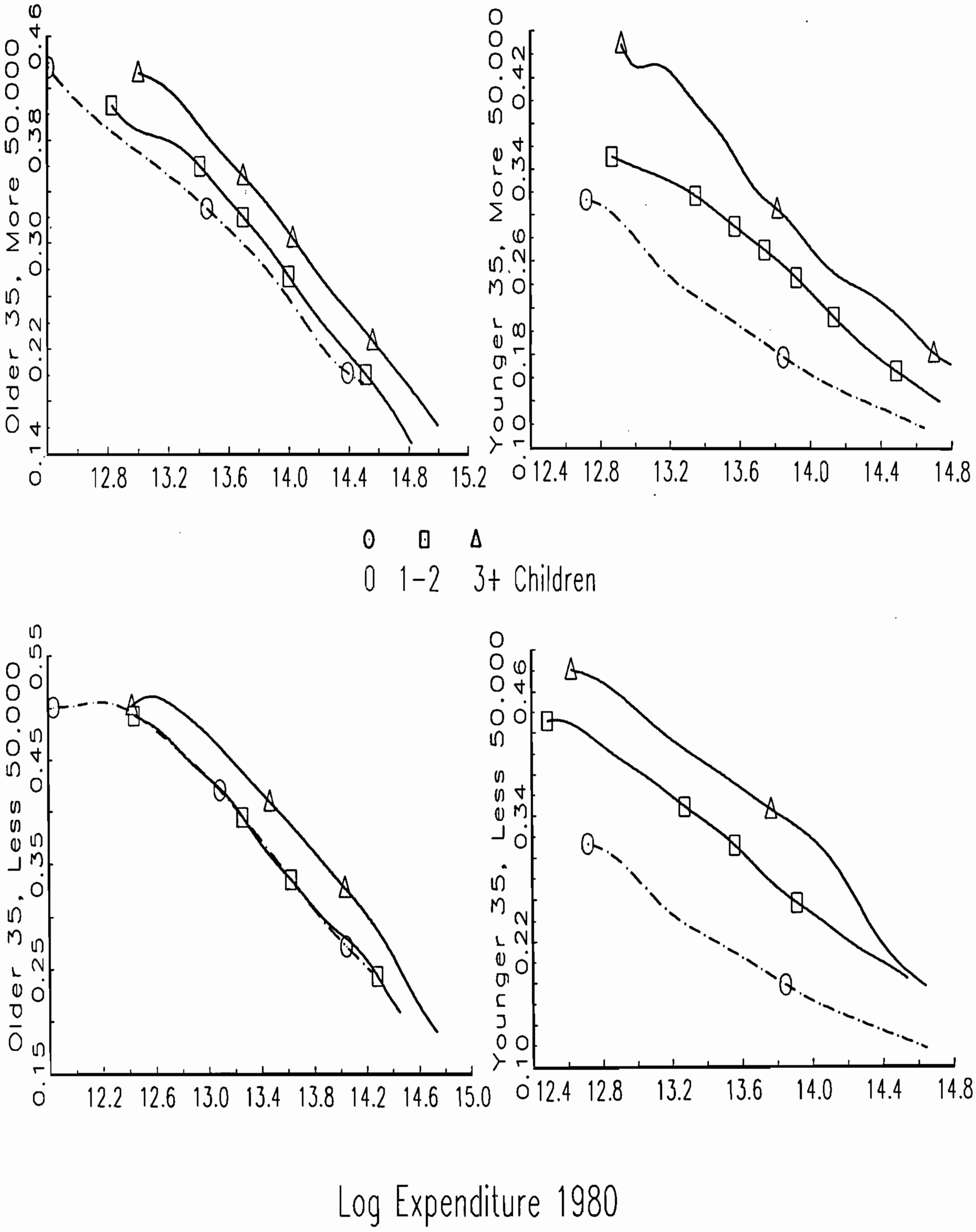


FIGURE 8

Nonparametric Estimate of Food Budget Share