Tilburg University

Measuring the costs of children

Melenberg, B.; van Soest, A.H.O.

Published in:
Statistica Neerlandica

Publication date:
1996

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):
Melenberg, B., & van Soest, A. H. O. (1996). Measuring the costs of children: Parametric and semiparametric estimators. *Statistica Neerlandica*, 50(1), 171-192.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Measuring the costs of children: parametric and semiparametric estimators¹

B. Melenberg* and A. H. O. van Soest

*Tilburg University, Dept. of Econometrics, P.O. Box 90153,
5000 LE Tilburg, The Netherlands*

How much additional income does a family with four children need to attain the same welfare level as a family with two children? And how much does a single person need compared to a childless couple? These questions are important in applied micro-economics and underlie public policy on, for example, social benefits and child allowances. In micro-economics, this is known as the problem of measuring equivalence scales. We estimate these scales using two types of subjective survey information. First, we use answers to survey questions on the income required to attain a given utility level. We compare the results for the usual linear model with semiparametric estimates, in which the functional form of the relationship between required income and family size and actual income is left unspecified. Second, we use answers to the question: how satisfied are you with actual household income? We present parametric and semiparametric estimates for ordered response models explaining this discrete variable. We find that according to the second type of information, costs of children are much larger than according to the first.

Key Words & Phrases: costs of children, equivalence scales, non- and semiparametric estimation, subjective information.

1 Introduction

A (household) equivalence scale is defined as the ratio between incomes needed to attain a given utility or welfare level for the family of interest and some reference family. Compare, for example, a family with two children (family 0) and a family with four children (family 1). If both families have the same income, the welfare level of family 0 is expected to be higher, since family 1 has to support more children. If the additional income required by family 1 to attain the welfare level of family 0 is 20 percent of income of family 0, the equivalence scale is 1.2.

¹ We are grateful to Marcel Das for his help and to Arie Kapteyn, participants at the HCM conference on Microeconomics of Public Policy in Leuven, and two referees for useful comments. Statistics Netherlands (CBS) has kindly provided the data. Financial support by the Royal Netherlands Academy of Arts and Sciences (KNAW) is gratefully acknowledged by the second author.

* b.melenberg@kub.nl

Equivalence scales are an important tool for applied welfare analysis and public policy (see, for example, BROWNING, 1992, NELSON, 1993, LEWBEL, 1989). They are used to define the costs of children, which serve as the basis for the system of child benefits in many countries. They are needed to define social minimum income levels for, for example, singles and married couples, and thus directly affect the minimum levels of social welfare benefits. Finally, they are indispensable for thorough analysis of income inequality within and between countries (see JENKINS, 1991): if household incomes are compared, equivalence scales are used to correct for differences in household composition.

By definition, equivalence scales describe disaggregate phenomena at the micro-level of the household. It is virtually impossible to estimate them using aggregate data. During the past decades, they have been a recurring issue in micro-econometric studies, with a large contribution of Dutch researchers. The traditionally most common approach to estimate them is via a demand system. Variation in expenditures on, for example, food and luxury goods across families with different composition, is then used to estimate the equivalence scales in an indirect way (see BROWNING, 1992).

Since POLLAK and WALES (1979), however, it has become well-known that expenditure data alone are not sufficient to identify the equivalence scales. See also KAPTEYN (1994) for a clear example illustrating this point, a discussion of possible solutions, and further references. Results of BLUNDELL and LEWBEL (1991) imply that the informational content of demand systems about equivalence scales is limited: for one set of prices, any equivalence scale is consistent with observed demand. On the other hand, relative changes in equivalence scales due to price changes are completely identified from demand data. For applied purposes, the latter is insufficient without the former. In section 2, we briefly discuss the economic framework and the identification issue.

The results of BLUNDELL and LEWBEL (1991) suggest that estimating equivalence scales could proceed in two steps. First, use other than demand data to estimate the levels of the equivalence scales in a given reference price setting. Second, use information on demand data to identify the effects of price changes on the equivalence scales. An enormous literature is devoted to this second step (see BROWNING, 1992). This paper is concerned with the first step only. We analyze equivalence scales in a fixed price setting. We avoid the complications and specification choices needed in order to formulate a demand system.

There are two types of information that we consider. The first has been used before by many, mainly Dutch researchers. See VAN PRAAG (1968, 1991), KAPTEYN (1994), and their references. It is based upon the following question

Which annual after-tax household income would you, in your circumstances, consider a) very bad; b) bad; c) insufficient; d) sufficient; e) good; f) very good?

The answers are an increasing series of six required income levels. We briefly discuss the literature using this question in section 3.

The second type of information is related to the first type: it comes from a survey question on satisfaction with actual family income. This type of information has been used by, for example, VAUGHAN (1984) and POULIN (1988). In section 4, we discuss the survey questions and the resulting sample data. We use a cross-section of Dutch households drawn from the Socio-Economic Panel (SEP), October 1984, which contains both types of information for the same households.

In sections 5 and 6, we discuss parametric and semiparametric models based upon the two types of subjective information. We focus on the equivalence scales, and try to avoid assumptions about functional form, distributions of error terms, etc. A flexible non- or semiparametric analysis therefore seems natural. We present and compare the results for various models. In the concluding section, we check whether the two types of information are compatible. We try to explain our finding that they are not, and argue why we think the second type of information is superior to the first.

Readers with little economic background may find it hard to read sections 2 and 3. Section 2 explains why we do not follow the traditional approach and use subjective data instead. Section 3 explains how our specification relates to others using similar data. Readers who are only interested in the statistical application can skip these sections.

2 Equivalence scales and the neoclassical demand system

We consider a household (indicated by index i) maximizing utility $u_i = u_i(q, h_i)$ under a linear budget constraint $p'q = y_i$. Here q denotes the vector of quantities of the goods involved (the choice variable of the household), p is the corresponding price vector, y_i is household income, and h_i is household composition. We assume that u_i satisfies the usual conditions of monotonicity and quasi-concavity. Utility may depend on i through h_i and through other variables not included in h_i . Let $c_i(u, h_i, p)$ be the household cost function, representing minimum expenditure required to attain a given utility level u . c_i is convex and increasing in p and strictly increasing in u (cf., for example, BARTEN and BÖHM, 1982). Let 0 denote the reference household. The definition of the equivalence scale (e) is

$$e_i(h_i, u, p) = c_i(u, h_i, p) / c_i(u, h_0, p)$$

or, in logs,

$$\ln c_i(u, h_i, p) = \ln c_i(u, h_0, p) + \ln e_i(h_i, u, p) \quad (1)$$

It says how large income of household i must be compared with income of household 0, if both households should attain the same utility level u . In general, the equivalence scale depends on the utility level. The hypothesis that it does not is often referred to as the independence of base utility (IB) assumption.

To obtain equations that can be estimated using demand data, Marshallian demand equations are used:

$$d_i(y_i, h_i, p) = \partial c_i(u, h_i, p) / \partial p, \text{ with } u \text{ determined by } c_i(u, h_i, p) = y_i \quad (2)$$

Here $d_i(y_i, h_i, p)$ is the vector of quantities that maximizes utility for household i , for given prices p and household income y_i .

Preferences and the utility function are completely characterized by the cost function. The source of the identification problem is that the cost function is not completely characterized by the Marshallian demand functions. The intuition behind this is clear: optimal quantities are those that maximize utility; a monotonic transformation of the utility function does not change the optimal quantities, while it does change the utility function and the cost function. To be precise: replace the cost function c_i by

$$C_i(u, h_i, p) = c_i(F(u, h_i), h_i, p)$$

where $F(u, h_i)$ is monotonically increasing in u . Then the demand functions corresponding to C_i are identical to those corresponding to c_i . The equivalence scales corresponding to C_i , however, are given by

$$C_i(u, h_i, p) / C_i(u, h_0, p) = c_i(F(u, h_i), h_i, p) / c_i(F(u, h_0), h_0, p)$$

Even if the original equivalence scales e_i do not depend on u , these equivalence scales will be different from the original ones if F depends on h_i . This shows that different equivalence scales can be consistent with the same demand system. By choosing F appropriately, BLUNDELL and LEWBEL (1991) show that, for fixed prices p , every set of equivalence scales is consistent with a given demand system. Thus demands for goods contain no information on equivalence scales in a single price system. Blundell and Lewbel also show that, once the equivalence scales in one price system p^0 are given, the demand functions completely identify the ratios $e_i(h_i, u, p) / e_i(h_i, u, p^0)$.

BLUNDELL and LEWBEL (1991) suggest several solutions to the identification problem. The first is not to estimate equivalence scales in levels, and look at ratios between equivalence scales for different price settings only. From a policy point of view, however, we are interested in levels and not in ratios. Levels are needed for applied welfare analysis concerning poverty, child benefits, income inequality, etc.

The second solution is to make additional assumptions on the utility functions, based on 'prior beliefs'. An example is the IB assumption: if IB is imposed, the demand system determines the equivalence scale (BLUNDELL and LEWBEL, 1991, p. 56). The IB assumption implies testable restrictions on the demand system, which have often been rejected. Other assumptions have been used, but, as the IB assumption, are to a large extent arbitrary and ad hoc (see, for example, NELSON, 1993).

We conclude that the only feasible solution is the third one: use additional data. This is the solution used by a school of researchers, primarily Dutch. We discuss it in the next section.

3 Using subjective data on income evaluation: earlier studies

This line of research starts with VAN PRAAG (1968) and has produced numerous publications by Van Praag, Kapteyn, and others. An early overview is given by KAPTEYN and WANSBEEK (1985). VAN PRAAG (1991) and KAPTEYN (1994) give more recent references. Most of the empirical analysis focuses on the question already mentioned in the introduction (usually referred to as the Income Evaluation Question (IEQ)):

Which annual after-tax household income would you, in your circumstances, consider a) very bad; b) bad; c) insufficient; d) sufficient; e) good; f) very good?

Across surveys, the number of incomes in the series varies, but the concepts are the same. The answers to this sort of question have been used to estimate equivalence scales in various ways. Many of these are variants on VAN PRAAG (1968), who assumes that the household cost function can be written as follows (since p is fixed, the argument p is omitted from now on):

$$\ln c_i(u, h_i) = \sigma_i(h_i)\Phi^{-1}(u) + \mu_i(h_i) \quad (3)$$

Φ is the standard normal distribution function. It is assumed that the utility level u depends on the question (a), b), ... or f)) but not on the household. In the case given above, the usual assumption is that $u = 1/12$, $u = 3/12$, ..., $u = 11/12$, for questions a) through f), respectively. Given the values of u and the observed values of $c_i(u, h_i)$, $\mu_i(h_i)$ and $\sigma_i(h_i)$ can be estimated for each household. Numerous articles have been written to model $\mu_i(h_i)$ as a function of household composition and other characteristics. The theory of relative perception of income and poverty (KAPTEYN, 1977) implies that $\mu_i(h_i)$ depends on actual income y_i . Empirical evidence strongly suggests that this is indeed the case. This can easily be incorporated in the framework of section 2, although it does not correspond directly to the neoclassical view.

Results of ALESSIE et al. (1995) suggest that $\mu_i(h_i)$ depends on demographics, habit formation, and incomes of families in the household's reference group. Recent results of MANSKI (1993) suggest, however, that nonparametric identification of reference group effects can only be obtained in very special circumstances: prior information on the composition of reference groups must be available, and the variables defining the reference groups and those directly affecting the outcomes should be 'moderately related'. Less attention has been paid to modelling $\sigma_i(h_i)$. Note that the IB assumption in the previous section is equivalent to the assumption that $\sigma_i(h_i)$ does not vary with h_i .

In this paper, we focus on equivalence scales. Habit formation, preference interdependence, etc. are not taken into account. We allow the perceived cost function to vary with actual household income. In (3), an explicit relationship is given between costs and utility, and it is implicitly assumed that utility is cardinal. This is something we also want to avoid. The only assumption we make is that, if different households answer the same question on which income is sufficient or very good, etc., they refer

to the same utility level. Starting point, therefore, is the equation of VAN PRAAG and VAN DER SAR (1988):

$$\ln c_i(u, h_i) = x_i' \beta(u) + H_i(u, h_i) + \gamma(u) \ln y_i + \varepsilon_i(u) \quad (4)$$

$c_i(u, h_i)$ is the observed answer to the question which income is sufficient, very good, etc., the required income to attain a given utility level. We consider the answers to the six questions separately and make no assumptions on how $c_i(u, h_i)$ depends on u . x_i is a vector of exogenous variables not related to family composition. We allow, for example, for regional variation (see section 5). $H_i(u, h_i)$ determines the equivalence scale for household i . In Van Praag and Van der Sar h_i simply represents family size. They use

$$H_i(u, h_i) = \alpha(u) \ln h_i \quad (5)$$

A more sophisticated specification is used by KAPTEYN et al. (1988):

$$H_i(u, h_i) = \alpha_0(u) \sum_j w_j f(a_j, u) \quad (6)$$

The sum is over all household members. The weights w_j are given by $w_1 = 1$ and $w_j = \ln(j/(j-1))$ for $j > 1$. $f(a_j, u)$ is the following function of the age a_j of individual j , where the ordering (within a household) is such that $a_j \geq a_{j+1}$:

$$\begin{aligned} f(a_j, u) &= 1 \text{ if } a_j > 18 \\ f(a_j, u) &= 1 + \alpha_1(u)(18 - a_j)^2 + \alpha_2(u)(18 - a_j)^2(36 + a_j) \text{ if } 0 \leq a_j \leq 18 \end{aligned} \quad (7)$$

Thus, $f(18, u) = 1$ and $f'(18, u) = 0$. α_0 , α_1 and α_2 are parameters to be estimated. By this specification the costs of children are allowed to be different from the costs of adults. The costs of children may also vary with age.

In (4), $H_i(u, h_i)$ determines the equivalence scale conditional upon income y_i . We are interested in unconditional equivalence scales, however. To eliminate $\ln y_i$, we again follow VAN PRAAG and VAN DER SAR (1988), and compute the so-called 'true' equivalence scale. This is the answer to the question: if we want two families to perceive the same utility level u , what is the ratio of the incomes the two families must have? We assume that other characteristics x_i and error terms (unobserved heterogeneity or measurement error) ε_i are the same for the two families. The income $y_i(u)$ necessary to reach utility level u can be obtained from (4), by setting $c_i(u, h_i)$ equal to $y_i = y_i(u)$:

$$\ln y_i(u) = (1 - \gamma(u))^{-1} (x_i' \beta(u) + H_i(u, h_i) + \varepsilon_i(u)) \quad (8)$$

Thus the log of the 'true' equivalence scale at utility level u for family i compared to the reference family 0 is given by:

$$\ln(y_i(u)/y_0(u)) = (1 - \gamma(u))^{-1} (H_i(u, h_i) - H_0(u, h_0)) \quad (9)$$

VAN PRAAG and VANDER SAR (1988) present results for many countries. Most of their data, including those for the Netherlands, are drawn from a 1979 European mail questionnaire. They find significantly positive estimates of $\gamma(u)$ in all cases. For the Netherlands, the values of $\gamma(u)$ vary from 0.42 to 0.63. The true equivalence scales they find hardly vary with the utility level. They are rather flat, i.e. the estimated costs of children are low. For example, compared to the reference case of a childless couple, they find equivalence scales of about 1.22 and about 1.37, for households with two and four children, respectively. The additional costs of the first two children would thus be about 22% of the childless couple's budget, while two more children would cost 12% of the budget of the reference family.

KAPTEYN et al. (1988) used (6) and (7) to estimate various models, using data from the Dutch labour mobility survey drawn in 1982. Their favourite model accounts for selection bias due to non-reporting, and incorporates that heads of household systematically underestimate certain components of their income (for example, family allowances and earnings of children). It also includes reference group averages. Although not presented, equivalence scales can be derived immediately from their results on poverty lines. For example, they find equivalence scales of 1.20 for a family with two children aged 6 and 12, and 1.34 for a family with four children, of age 18, 12, 6 and 1 (both compared with a childless couple). KAPTEYN and MELENBERG (1993), who apply a similar specification to the SEP-data (wave 1985), find still lower costs of children.

4 Data

Data are taken from the October 1984 wave of the Dutch Socio-Economic Panel (SEP), drawn by Statistics Netherlands. The SEP is fairly representative for the Dutch population (see CBS, 1991). The original data set contains about 3700 households, with accurate information on income from forty potential sources for each individual. After tax household income was constructed by adding up all income components of all family members. Missing observations, for example, for self-employed household members, led to a loss of about 1000 families. Some households were excluded because of missing information on family composition or other explanatory variables.

We consider two types of subjective information. The first is required income to attain some utility level (see section 1). We denote the utility levels corresponding to very bad, bad, insufficient, sufficient, good, and very good, by 1, 2, 3, 4, 5 and 6, respectively. 2455 families answer all these questions. Required family income increases with utility level in all cases. In Figure 1, we present nonparametric (Gaussian kernel) density estimates of required and actual monthly incomes. Required incomes for different utility levels have similar distributions. Dispersion increases with utility level. Actual incomes reveal a much larger dispersion. The location of the distribution of actual incomes compared to that of required incomes, suggests that most families are reasonably satisfied with their income.

Fig. 1. Distribution of actual income and of required income for utility levels 1, ..., 6.

Fig. 2. Distribution of income required for utility level 3.

In Figure 2, we have drawn the distribution of required income for utility level 3 with varying household size. Required income for singles is much smaller than for other households. The differences between families with 2, 3, 4 or more members, are surprisingly small. For other utility levels, the figure is similar. In Figure 3, we present the results of a nonparametric (quartic kernel) regression of the required income

Fig. 3. Nonparametric regression of required income (utility levels 1, ..., 6) on actual income, with uniform confidence bands.

levels on family income (not controlling for family size). We also present 95% uniform confidence bounds for the regression function, based upon HÄRDLE and LINTON (1994). These become very wide where observations are sparse. The figure suggests that a linear relationship between actual and required income might be appropriate for the range where the bounds are tight, but not for the complete range of incomes in the sample. For the range of very low incomes, we find that required

Fig. 4. Nonparametric regression of satisfaction with actual income (SAI) on actual income, with uniform 95% confidence bands.

income does not increase with actual income. This is contrary to what we would expect. It might be explained by the relatively low number of observations in this income range.

The second type of subjective information is the answer given by the head of the household to the question

How satisfied are you with your household income? Possible answers 1 (not satisfied at all) to 10 (very satisfied).

2729 heads of household have answered this question. The variable is denoted by *SAI*. In 772 cases, *SAI* is less than or equal to 5. The sample average is 6.30. The dispersion is substantial. 74 observations have *SAI* = 1; 101 have *SAI* = 10. In Figure 4, we present the results of a nonparametric (quartic kernel) regression of *SAI* on family income. As expected, satisfaction with income increases with income. It appears to decrease with family size. In section 6, we analyze what this means for the equivalence scales.

5 Empirical results: required income questions

We first analyze the required income questions. We considered a number of specifications, starting with equations (4) and (5) in section 3:

$$\ln c_i(u, h_i) = x_i' \beta(u) + \alpha(u) \ln h_i + \gamma(u) \ln y_i + \varepsilon_i(u) \quad (10)$$

$c_i(u, h_i)$ is required income to attain a given fixed utility level u for a family with size h_i . It is allowed to depend on actual income y_i , since people with higher income will often feel they need more. Estimated 'true' equivalence scales comparing families 1 and 0 are obtained by setting required and actual income equal to each other and solving for income, assuming exogenous variables and error terms are equal for the two families. Then ratios of the incomes of the two families are computed (see (8) and (9)):

$$\ln(y_1(u)/y_0(u)) = (1 - \gamma(u))^{-1} \alpha(u) \ln(h_1/h_0) \quad (11)$$

They can be estimated using the estimates of (10). Results are presented in Table 1. To save space, (OLS) estimation results of (10) are not presented. Gender of the head of the household, regional dummies and the degree of urbanization are included in x_i in all the regressions. Some of these are strongly significant. In the first panel of Table 1, we present the results for the specification most similar to that of Van Praag and Van der Sar ((4)–(5)). We find lower costs of children than they do. We also find larger (and sometimes significant) differences between the utility levels. Costs of children decrease with the utility level. The small standard errors suggest that the results are quite accurate.

To allow for the nonlinearities suggested by Figure 3, we added log family size squared, log actual income squared, and the product of log income and log family size. The latter two appeared to be strongly significant for all utility levels. The resulting true equivalence scales are in the second panel of Table 1. Note that equation (11) could no longer be used here: true equivalence scales had to be obtained from solving a quadratic equation. The solution always existed. Costs of children are now strongly decreasing with the utility level. For the highest utility level, the estimated costs of children are (insignificantly) negative. This seems implausible, although it might be interpreted if children themselves are seen as consumption goods. For the lower utility levels, the costs of the first child are larger than according to the linear specification, while the costs of more than one child are similar. Standard errors have approximately doubled.

Differences between the two panels suggest that we should consider more flexible specifications. We turn to a semiparametric model, in which the impact of family size and actual income on required income is left unspecified. Equation (10) is replaced by

$$\ln c_i(u, h_i) = x_i' \beta(u) + g_u(h_i, \ln y_i) + \varepsilon_i(u) \quad (12)$$

Here g_u is a continuous function satisfying some regularity conditions. As before, we assume $E(\varepsilon_i(u) | x_i, h_i, y_i) = 0$. Equation (12) leads to the partial regression model of ROBINSON (1988) and STOCK (1991). The latter proposes a three step estimator for this model (with unknown asymptotic efficiency characteristics), based upon the following identity, derived from (12):

$$\ln c_i(u, h_i) - E(\ln c_i(u, h_i) | h_i, \ln y_i) = (x_i - E(x_i | h_i, \ln y_i))' \beta(u) + \varepsilon_i(u) \quad (13)$$

Table 1a. 'True' equivalence scales Van Praag & Van der Sar model

FS	C1	C2	C3	C4	C5	C6
1	0.895	0.907	0.912	0.917	0.931	0.953
	0.010	0.010	0.011	0.011	0.013	0.016
2	1.000	1.000	1.000	1.000	1.000	1.000
	0.000	0.000	0.000	0.000	0.000	0.000
3	1.067	1.059	1.055	1.052	1.043	1.028
	0.007	0.007	0.007	0.007	0.008	0.010
4	1.117	1.102	1.096	1.091	1.074	1.049
	0.013	0.013	0.013	0.013	0.014	0.018
6	1.192	1.167	1.157	1.148	1.121	1.079
	0.021	0.021	0.022	0.022	0.024	0.029

Table 1b. 'True' equivalence scales extended version Van Praag & Van der Sar (with squared and cross terms)

FS	C1	C2	C3	C4	C5	C6
1	0.799	0.820	0.837	0.872	0.918	1.041
	0.020	0.019	0.019	0.021	0.028	0.074
2	1.000	1.000	1.000	1.000	1.000	1.000
	0.000	0.000	0.000	0.000	0.000	0.000
3	1.095	1.087	1.081	1.064	1.037	0.985
	0.012	0.011	0.010	0.009	0.009	0.016
4	1.150	1.139	1.131	1.105	1.059	0.977
	0.024	0.022	0.020	0.019	0.019	0.026
6	1.207	1.198	1.192	1.156	1.085	0.968
	0.049	0.044	0.043	0.039	0.039	0.047

Table 1c. 'True' equivalence scales semiparametric model

FS	C1	C2	C3	C4	C5	C6
1	0.890	0.904	0.918	0.941	0.969	1.006
	0.031	0.028	0.028	0.030	0.036	0.052
2	1.000	1.000	1.000	1.000	1.000	1.000
	0.000	0.000	0.000	0.000	0.000	0.000
3	1.050	1.043	1.037	1.026	1.012	0.993
	0.032	0.026	0.024	0.024	0.027	0.035
4	1.067	1.057	1.048	1.033	1.012	0.983
	0.034	0.028	0.025	0.025	0.027	0.034
6	1.073	1.060	1.050	1.033	1.005	0.964
	0.045	0.035	0.030	0.028	0.030	0.036

Explanation: FS=Family size; C1: corresponds to level 1; C2: corresponds to level 2, and so on. Each first row gives the estimated equivalence scales; Each second row gives the estimated standard error.

The first step consists of replacing the expectations in (13) by their nonparametric (kernel) regression estimates. The next step is to estimate β by OLS on (12). The third step is to estimate g_u by nonparametric regression of $\ln c_i(u, h_i) - x_i' \hat{\beta}(u)$ on h_i and $\ln y_i$, where $\hat{\beta}(u)$ is the OLS estimate of $\beta(u)$. Finally, the 'true' equivalence scales can be computed by (numerically) solving $y_i(u, h)$ from

$$\ln y_i(u, h) = x_i' \beta(u) + \hat{g}_u(h_i, \ln y_i(u, h)) \quad (14)$$

Table 2. 'True' equivalence scales with age function Kapteyn et al. (1988) included

NA	K1	K2	K3	K4	K5	K6	no cross terms		cross terms	
							C1	St err	C1	St err
1	—	—	—	—	—	—	0.913	0.015	0.833	0.028
2	—	—	—	—	—	—	1.000	0.000	1.000	0.000
2	6	—	—	—	—	—	1.086	0.012	1.047	0.038
2	12	—	—	—	—	—	1.067	0.008	1.069	0.022
2	12	6	—	—	—	—	1.131	0.016	1.106	0.047
2	18	12	6	1	—	—	1.200	0.021	1.229	0.038
NA	K1	K2	K3	K4	K5	K6	C3	St err	C3	St err
1	—	—	—	—	—	—	0.927	0.017	0.886	0.026
2	—	—	—	—	—	—	1.000	0.000	1.000	0.000
2	6	—	—	—	—	—	1.073	0.012	1.054	0.032
2	12	—	—	—	—	—	1.060	0.008	1.057	0.018
2	12	6	—	—	—	—	1.114	0.015	1.099	0.038
2	18	12	6	1	—	—	1.159	0.021	1.176	0.030
NA	K1	K2	K3	K4	K5	K6	C6	St err	C6	St err
1	—	—	—	—	—	—	0.932	0.018	0.926	0.020
2	—	—	—	—	—	—	1.000	0.000	1.000	0.000
2	6	—	—	—	—	—	1.048	0.013	1.044	0.015
2	12	—	—	—	—	—	1.048	0.009	1.047	0.010
2	12	6	—	—	—	—	1.083	0.017	1.080	0.018
2	18	12	6	1	—	—	1.118	0.023	1.113	0.022

Explanation: NA=Number of adults; K1–K6: Age of the children; — means: not included; C1 corresponds to level 1, C3 to level 3, C6 to level 6. St err: Estimated standard error.

and computing the ratio between $y_i(u, h)$ for the family of interest and the reference family with the same value of x_i . Results are in the third panel of Table 1. Following STOCK (1991), standard errors can be computed without taking account of the first two steps, due to differences in the rates of convergence. From pointwise asymptotic properties of $\hat{g}_u$ (cf. BIERENS, 1987), it is then possible to compute the standard errors of the equivalence scales. Details are available upon request from the authors. Again, we find that the costs of children would be quite low, confirming the results obtained by the parametric second order specification. Only the difference between a single person and a childless couple is significant (for the lower utility levels), costs of children are insignificant in all cases. This suggests that costs of children could well be lower than according to the linear specification, which is rejected by the data.

To allow the equivalence scales to vary with the children's ages, we have also estimated some specifications using (6) and (7) instead of (5). The impact of household composition h_i on the costs of children then not only depends on family size, but also on the ages of the children. A summary of the results is presented in Table 2. For the linear specification, the results correspond to those in Table 1. Again, the costs of children are quite low, even lower than those found by KAPTEYN and MELENBERG (1993). Age of children does not play a large role. In many cases, estimated costs of young children are higher than those of older children. Adding cross terms and squares leads to some changes (second panel of Table 2). The costs of

a child are now increasing with the age of the child. The estimated costs of children remain low. While the costs of young children correspond to findings of KAPTEYN et al. (1988), those of older children and adults are much lower. The standard errors suggest that these results are reasonably accurate. We did not estimate the semi-parametric generalization of this specification, because our data set is not large enough to tackle the higher dimensional estimation problem involved with this.

6 Satisfaction with actual income: models and results

The dependent variable SAI is discrete with possible values 1 through 10. An ordered response model is therefore appropriate:

$$y_i^* = z_i' \gamma + \varepsilon_i; \quad SAI_i = j \text{ if } m_{j-1} < y_i^* < m_j \quad (j = 1, \dots, 10) \quad (15)$$

y_i^* is a latent variable. z_i is a vector of regressors including x_i defined in section 5, family composition, and log family income. The bounds m_j satisfy $-\infty = m_0 < m_1 < \dots < m_9 < m_{10} = \infty$. By means of normalization, we choose $m_1 = 1.5$ and $m_9 = 9.5$. The other m_j are estimated. (The hypothesis that $m_j = j + 1/2$ for $j = 2, \dots, 8$, is rejected for all models.) ε_i is the error term. The equivalence scales can easily be derived from (15): the log equivalence scale is equal to the log family income difference needed to compensate for a change in family composition. If log family income enters linearly, this is simply the ratio of the family composition effect and the coefficient of log family income. Standard practice is to assume

$$\varepsilon_i | z_i \sim N(0, \sigma^2) \quad (16)$$

This yields the ordered probit model (see MADDALA, 1983), which can easily be estimated by maximum likelihood (ML). However, ML may be inconsistent if (16) is not satisfied. Score tests (as in CHESHER and IRISH, 1987) suggest that nonnormality or heteroskedasticity of the errors lead to violation of (16). We therefore also consider some more flexible parametric and semiparametric specifications.

A parametric generalization of (16) is given by

$$P(\varepsilon_i \leq t | z_i) = G(t / \exp(z_i' \alpha)), \text{ with } G(s) = \Phi(s + \psi_1 s^2 + \psi_2 s^3) \quad (17)$$

Here Φ is the standard normal distribution function. (17) allows for (exponential) heteroskedasticity and nonnormality. Normality is obtained for $\psi_1 = \psi_2 = 0$. For arbitrary ψ_1 and ψ_2 , we get the family of distributions introduced by RUUD (1984). Monotonicity of G implies that $\psi_1^2 - 3\psi_2$ must be nonpositive.

Semi-nonparametric estimation (Gallant/Nychka)

A more flexible family of nonnormal distributions is introduced by GALLANT and NYCHKA (1987) (G&N, from now on). We replace (16) by

$$\varepsilon_i | z_i \sim h \in H \quad (18)$$

Here H is some class of univariate densities satisfying 'mild regularity conditions' (G&N, p. 387). Ruled out are densities with too violent oscillations or too fat or too thin tails. Skewness etc. is permitted. The main idea is that any density in H can be approximated by an element in the set H_K of densities of the form

$$h(u) = (\sum_{k=0}^K \alpha_k u^k)^2 f(u, \sigma), \quad \text{with } f(u, \sigma) = (2\pi\sigma)^{-1/2} \exp(-u^2)/(2\sigma^2) \quad (19)$$

The vector $\alpha \in \mathbb{R}^{K+1}$ must be such that the density is proper (see below). The approximation can be made arbitrarily close (in some norm on H) if K is large enough, i.e. the union of all H_K ($K = 1, 2, \dots$) is dense in the closure of H . This implies that, under additional regularity conditions guaranteeing uniform convergence and identification, a consistent estimator for γ can be found by maximizing the likelihood over γ , $m_2, \dots, m_8$ and α , provided that $K \rightarrow \infty$ with the number of observations N (Theorem 0 in G&N). The motivation for the G&N densities is thus based on their properties for large K . In practice, K is chosen fixed and finite. If K is fixed, the model is parametric and usual ML properties are valid, under the usual assumption that H_K contains the true density.

A first advantage of (19) compared to other series approximations is that computing the probabilities $P(SAI_i = j | z_i)$ is easy, using higher order truncated moments of the univariate normal distribution. Second, the density is guaranteed to be nonnegative. Third, this family nests the univariate normal distribution (with zero mean). The number of terms to be computed for the likelihood contribution per observation, increases with K^2 .

Some normalizations have to be added for practical implementation. First, the density must be proper: $\int_{\mathbb{R}} h(u) du = 1$. This implies a restriction on α . We impose this in the same way as GABLER et al. (1993): start with an improper density h^* determined by α^* , with $\alpha_0^* = 1$. Compute the integral for α^* , and divide the α_k^* by its square root to obtain α . Second, nonparametric identification of the model requires a location restriction on the errors or the systematic part. We a priori fix the constant term to its probit estimate. This is easier than imposing restrictions on α to guarantee that the error has mean zero, the procedure followed by GABLER et al. (1993).

Semiparametric estimation (Horowitz/Lee)

The semi-nonparametric estimator described above maintains the assumption that errors and regressors are independent. It does not allow for heteroskedasticity. A more general model is obtained if (16) is replaced by

$$\text{Median}(\varepsilon_i | z_i) = 0 \quad (20)$$

LEE (1992) has proposed a consistent estimator for the model defined by (15) and (20), based upon the observation that (15) and (20) imply

$$\text{Median}(SAI_i | z_i) = 1 + \sum_{r=1}^9 1(z'_i \gamma \geq m_r) \quad (21)$$

where $1(\cdot)$ denotes the indicator function. (21) immediately follows from the median

preserving property of increasing functions. It implies

$$(\gamma, m_2, \dots, m_8) = \text{Argmin } E(|SAI_i - 1 - \sum_{r=1}^9 1(z'_i \gamma \geq m_r)|) \quad (22)$$

It follows that $(\gamma, m_2, \dots, m_8)$ can be estimated consistently by least absolute deviations:

$$(\hat{\gamma}, \hat{m}_2, \dots, \hat{m}_8) = \text{Argmin } (\sum_i |SAI_i - 1 - \sum_{r=1}^9 1(z'_i \gamma \geq m_r)|) \quad (23)$$

This is the estimator of LEE (1992). It generalizes the maximum score estimator of MANSKI (1985) for the binary choice model. It has the same drawback as Manski's estimator: the asymptotic distribution is intractable. For the maximum score estimator, this problem is solved by HOROWITZ (1992). His 'smoothed maximum score' estimator maximizes a smoothed version of the sum of least absolute deviations. The same idea can be applied here. The smoothed least absolute deviations estimator is obtained as follows (see MELENBERG and VAN SOEST, 1994, for details). Let $y_{ir} = 1(y_i^* \geq m_r)$. Then $y_i = \sum_r y_{ir}$. It is easy to verify that the minimization problem (23) can be rewritten as

$$(\hat{\gamma}, \hat{m}_2, \dots, \hat{m}_8) = \text{Argmax } \sum_i \sum_r (2y_{ir} - 1) 1(z'_i \gamma - m_r \geq 0) \quad (24)$$

Similar to HOROWITZ (1992), we can smooth $1(z'_i \gamma - m_r \geq 0)$ by $K((z'_i \gamma - m_r)/\sigma_N)$, where K is a function satisfying $\lim_{\nu \rightarrow \infty} K(\nu) = 1$ and $\lim_{\nu \rightarrow -\infty} K(\nu) = 0$, and where σ_N is a smoothness parameter depending on the sample size N , with $\lim_{n \rightarrow \infty} \sigma_N = 0$. Thus, the smoothed version of (24) becomes

$$(\hat{\gamma}, \hat{m}_2, \dots, \hat{m}_8) = \text{Argmax } \sum_i \sum_r (2y_{ir} - 1) K((z'_i \gamma - m_r)/\sigma_N) \quad (25)$$

This smoothed estimator is similar to Horowitz's smoothed Maximum Score estimator, and shares its asymptotic characteristics: consistency and the limit distribution. Convergence is slower than $N^{1/2}$. The convergence rate depends on the smoothness of the underlying model and the choice of the function K . Following HOROWITZ (1992), we use

$$\begin{aligned} K(\nu) = & (0.5 + (105/64)((\nu/5) - (5/3)(\nu/5)^3 \\ & + (7/5)(\nu/5)^5 - (3/7)(\nu/7)^7)) 1(-5 \leq \nu \leq 5) \end{aligned} \quad (26)$$

Under appropriate smoothness conditions, the limit distribution of the estimator defined by (25) then satisfies

$$N^{4/9}((\hat{\gamma}, \hat{m}_2, \dots, \hat{m}_8) - (\gamma, m_2, \dots, m_8)) \rightarrow N(AB, V) \quad (27)$$

Consistent estimates for the asymptotic bias AB and the asymptotic covariance matrix V can be constructed analogously to HOROWITZ (1992). See MELENBERG and VAN SOEST (1994). To solve (25), local search algorithms are infeasible, because the objective function is not concave and has several local maxima. We use the simulated

annealing type global search algorithm taken from CORANA et al. (1987), and discussed by GOFFE et al. (1994).

Partial linear model (Robinson/Stock)

The models introduced above have various degrees of flexibility concerning the conditional distribution of the error terms, but do not allow for nonlinearities in the systematic part. We consider one additional specification in which the systematic part is not assumed to be linear. Instead of using an ordered response model, we can directly model the mean of the observed variable SAI_i . We use a partial linear model, as in equation (12):

$$SAI_i = x_i' \beta + g(\ln y_i, fs_i) + \varepsilon_i; \quad E(\varepsilon_i | x_i) = 0 \quad (28)$$

g is an unknown function to be estimated. This model is estimated in the way described in section 5, following STOCK (1991). In this model, the true equivalence scales can be computed numerically from the estimates of g . Unlike in the linear specification, the outcome generally depends on the reference level of income, or, equivalently, utility. We compute them for a household with ε_i , x_i and utility level SAI_i set equal to their sample average. For each household of size h , we compute $\ln y$ from

$$\hat{g}(\ln y, fs) = 6.273 - \bar{x}' \hat{\beta}$$

We then take the ratios of required incomes for families of different composition.

Results

We first consider equivalence scales depending on family size only. The results are presented in Table 3. The probit (col. 1), extended probit (col. 2), and semi-nonparametric (col. 3; $K = 3$) estimates are similar. They imply much larger costs of children

Table 3. Equivalence scales based on Satisfaction with Actual Income (SAI)

FS	Pr	Ext Pr	Gall/N	Hor/L	Partial LM
1	0.553 0.024	0.585 0.022	0.556 0.024	0.659 0.012	0.726 0.117
2	1.000 0.000	1.000 0.000	1.000 0.000	1.000 0.000	1.000 0.000
3	1.414 0.036	1.369 0.030	1.410 0.036	1.277 0.014	1.280 0.205
4	1.807 0.079	1.711 0.065	1.799 0.079	1.518 0.028	1.485 0.255
6	2.554 0.177	2.342 1.141	2.535 0.177	1.938 0.056	1.838 0.414

Explanation: FS=Family size; Each first row gives the equivalence scale; Each second row gives the estimated standard error. Pr: Ordered Probit; Ext Pr: Extended version Ordered Probit; Gall/N: Gallant & Nychka; Hor/L: Horowitz & Lee; Partial LM: Semiparametric partial model.

than the results in the previous section. For example, the costs of two children would amount to more than 70% of the budget of a childless couple.

The semiparametric Horowitz/Lee estimator leads to smaller estimates of the costs of children. For example, the costs of two children would amount to about 52% of the childless couple's budget (standard error: 3%). This is still much more than in section 5 or in the related literature. The standard errors of this estimator are much lower than those of the other models. This finding is similar to what other researchers have found when applying the Horowitz estimator to a binary choice model. According to HOROWITZ (1992), the low standard errors indicate that the sample size is not yet large enough to make the asymptotic results applicable.

The partial linear model estimates are largely in line with the semiparametric estimates based upon the ordered response model. The standard errors, however, are huge. For example, the estimated costs of the first two children are 58% of the childless couple's budget, with corresponding standard error 35%. The estimates look quite plausible, but, according to the standard errors, could be very inaccurate. Similar results were obtained for other values of the smoothness parameters involved in the nonparametric kernel regression step.

In Table 4, we present results of the ordered response models in which costs of children are allowed to vary with the children's age levels, as in (6) and (7). The probit model leads to very large costs of children. Its parametric extension, which is a significant improvement considering the likelihood ratio, results in somewhat smaller costs of children. Similar outcomes are obtained using the semi-nonparametric model. Lower costs of children are obtained using the semiparametric Horowitz/Lee estimator. For example, the costs of the first two children amount to about 36% of the childless couple's budget. This is still much larger than the estimates in section 5.

Table 4. Equivalence scales based on Satisfaction with Actual Income (SAI), with age function Kapteyn et al. (1988) included

NA	K1	K2	K3	K4	K5	K6	Pr	St err	Ext Pr	St err
1	—	—	—	—	—	—	0.531	0.029	0.573	0.025
2	—	—	—	—	—	—	1.000	0.000	1.000	0.000
2	6	—	—	—	—	—	1.341	0.071	1.251	0.056
2	12	—	—	—	—	—	1.401	0.046	1.333	0.037
2	12	6	—	—	—	—	1.725	0.113	1.563	0.088
2	18	12	6	1	—	—	2.482	0.182	2.110	0.127
NA	K1	K2	K3	K4	K5	K6	Gall/N	St err	Hor/L	St err
1	—	—	—	—	—	—	0.558	0.024	0.581	0.013
2	—	—	—	—	—	—	1.000	0.000	1.000	0.000
2	6	—	—	—	—	—	1.276	0.056	1.119	0.021
2	12	—	—	—	—	—	1.355	0.036	1.258	0.015
2	12	6	—	—	—	—	1.611	0.087	1.362	0.033
2	18	12	6	1	—	—	2.215	0.125	1.842	0.053

Explanation: NA=Number of adults; K1–K6: Age of the children; — means: not included; Pr: Ordered Probit; Ext Pr: Extended version Ordered Probit; Gall/N: Gallant & Nychka; Hor/L: Horowitz & Lee; St err: Estimated asymptotic standard errors.

As before, the estimated standard errors of the semiparametric estimator are quite low. According to all specifications, the costs of a child increase with the child's age. The four estimators lead to substantially different estimates of the costs of a first child of six years old. We did not estimate a partial linear regression model based on specification (6) and (7), because of expected dimensionality problems and because of the large standard errors in Table 3.

7 Summary and concluding remarks

We have analyzed equivalence scales for Dutch households. We have first summarized the results in the literature on identification. This has led to the conclusion that estimating equivalence scales proceeds of two steps: first, estimate the equivalence scales for a given fixed set of prices; second, use demand data to analyze the consequences of price changes for the relative equivalence scales. Our analysis deals with the first step only. Consumer demand data are not helpful here.

Instead, other, often subjective, information, can be used. We have presented a brief overview of the mainly Dutch literature in this field, based upon the use of information on household income required to attain some given utility level. We have focused on studies that use as few additional assumptions as possible. This has led to the specification in which, for each utility level separately, required income is explained from family composition, the level of actual income, and background variables.

Using data from the Socio-Economic Panel, we have estimated a linear specification and various nonlinear specifications, both parametric and semiparametric. The main finding is that estimates of the costs of children based upon the required income data are implausibly low. For the more flexible specifications, this conclusion emerges stronger than for the linear specification.

We have then turned to another type of subjective information, concerning the satisfaction with actual current household income. We have shown that it can be used to derive equivalence scales in a very simple way. Several estimators of ordered response models are compared: ordered probit, a parametric extension, a semi-nonparametric extension, and a general semiparametric extension. The estimators differ with respect to the assumptions concerning the distribution of the error terms. We have also presented the results of a partial linear model, in which the functional form of the relation between income, family size, and expected satisfaction with income, is left unspecified. Compared to the estimates based upon the required income data, all these models lead to much higher costs of children. The accuracy of the results varies strongly with the chosen specification.

Because of the large variation of the standard errors obtained with the models explaining satisfaction with income, we did not formally test whether equivalence scales obtained with the two types of information are compatible. Still, in order to get some idea about the compatibility of the two types of information, we constructed a

second measure of satisfaction with actual income (*SAI2*) from the required income questions, as follows:

$$SAI2_i = 3/2 \sum_k 1(y_i \geq c_i(u_k, h_i)) + 1$$

The sum is taken over the six utility levels. If $c_i(u_k, h_i)$, required income as provided by the head of household, is lower than actual household income y_i , then we know that the household's utility level is at least u_k . This leads to an indicator of satisfaction with actual income, which has been transformed to a 1 to 10 scale, assuming that the $u_1, \dots, u_6$ correspond to 2.5, $\dots$, 8.5, and 10, respectively. For example, $SAI2_i = 8.5$ means that actual family income is between the incomes considered 'good' and 'very good'.

If the two types of information would be compatible, *SAI2* and *SAI* should be similar. Their rank correlation coefficient is only 0.32, however. For example, about 11% of all families have a value of *SAI* of at most 5, but a value of at least 8.5 for *SAI2*. The average value of *SAI2* is 7.79, while the average value of *SAI* is only 6.27. Apparently, the two measures of satisfaction with actual income are incompatible.

Several explanations for this finding are possible. The first is that our allocation of the six utility levels 'very bad', 'bad', $\dots$, 'very good' to the values of 2.5, 4, $\dots$, 10 on the *SAI* scale is wrong. This allocation is in line with the theory of cardinal utility functions developed by KAPTEYN (1977). A second, more plausible, interpretation is that our objective measure of actual family income does not correspond to the income concept used by the head of the household when answering the required income question. KAPTEYN et al. (1988) compare objective actual household income with categorical information on household income directly provided by the head of the household, and show that certain components of household income are strongly underestimated. For example, the head of the household has almost perfect knowledge of his own earnings, but tends to report less than 80% of the spouse's earnings, and less than 30% of earnings of children or family allowance. If the head of the household underestimates required income in a similar way as actual household income, this could explain the deviation between the two measures of satisfaction with actual income, as well as the low estimates of costs of children from the required income information.

The direct question about satisfaction with income does not require that the respondent estimates household income. We tend to have more confidence in the outcomes based upon this than in the outcomes based upon the required income questions. Ex ante, we would tend to believe that the satisfaction with actual income question is easier to answer than the required income questions. The former refers to the family's actual situation, while the latter refers to some virtual situation, which seems less well-defined. Ex post, we find the former outcomes more plausible than the latter.

References

- ALESSIE, R., A. KAPTEYN and B. MELENBERG (1995), Subjective poverty definitions and reference groups, mimeo, Tilburg University.
- BARTEN, A. and V. BÖHM (1982), Consumer Theory, in: K. Arrow and M. Intriligator (eds), *Handbook of mathematical economics*, Vol. II, North-Holland, Amsterdam.
- BIERENS, H. (1987), Kernel estimators of regression functions, in: T. F. Bewley (ed.) *Advances in econometrics, Fifth world congress*, Vol. I, Cambridge University Press, New York.
- BLUNDELL, R. and A. LEWBEL (1991), The information content of equivalence scales, *Journal of Econometrics* **50**, 49–68.
- BROWNING, M. (1992), Children and household economic behavior, *Journal of Economic Literature* **30**, 1434–1475.
- CBS (1991), Sociaal-economisch panelonderzoek, mimeo, Statistics Netherlands, Voorburg.
- CHESHER, A. and M. IRISH (1987), Residual analysis in the grouped and censored normal linear model, *Journal of Econometrics* **34**, 33–61.
- CORANA, A., M. MARCHESI, C. MARTINI and S. RIDELLA (1987), Minimizing multimodal functions of continuous variables with the 'simulated annealing algorithm', *ACM Transactions on Mathematical Software* **13**, 262–280.
- GABLER, S., F. LAISNEY and M. LECHNER (1993), Semi-nonparametric maximum likelihood estimation of binary choice models with an application to labour force participation', *Journal of Business and Economic Statistics* **11**, 61–81.
- GALLANT, A. R. and D. W. NYCHKA (1987), Semi-nonparametric maximum likelihood estimation, *Econometrica* **55**, 363–390.
- GOFFE, W. C., G. D. FERRIER and J. ROGERS (1994), Global optimization of statistical functions with simulated annealing, *Journal of Econometrics* **60**, 65–99.
- HÄRDLE, W. and O. LINTON (1994), Applied nonparametric methods, in: R. Engle and D. McFadden (eds.), *Handbook of econometrics*, vol. IV, Elsevier, Amsterdam.
- HOROWITZ, J. L. (1992), A smoothed maximum score estimator for the binary response model, *Econometrica* **60**, 505–531.
- JENKINS, S. P. (1991), The measurement of income inequality, in: L. Osberg (ed.), *Economic inequality and poverty*, M. E. Sharpe, New York.
- KAPTEYN, A. (1977), A theory of preference formation, PhD dissertation, Leyden University.
- KAPTEYN, A., P. KOOREMAN and R. WILLEMSE (1988), Some methodological issues in the implementation of subjective poverty definitions, *Journal of Human Resources* **23**, 222–242.
- KAPTEYN, A. and B. MELENBERG (1993), Measuring the well-being of the elderly, in: A. Atkinson and M. Rein (eds.), *Age, work and social security*, Macmillan, London.
- KAPTEYN, A. (1994), The measurement of household cost functions: revealed preference versus subjective measures, mimeo, Tilburg University.
- KAPTEYN, A. and T. WANSBEEK (1985), The individual welfare function, *Journal of Economic Psychology* **6**, 333–363.
- LEE, M.-J. (1992), Median regression for ordered discrete response, *Journal of Econometrics* **51**, 59–77.
- LEWBEL, A. (1989), Household equivalence scales and welfare comparisons, *Journal of Public Economics* **39**, 377–391.
- MADDALA, G. S. (1983), *Limited dependent and qualitative variables in econometrics*, Cambridge University Press, Cambridge (Mass.).
- MANSKI, C. (1985), Semiparametric analysis of discrete response, *Journal of Econometrics* **27**, 313–333.
- MANSKI, C. (1993), Identification of endogenous social effects: the reflection problem, *Review of Economic Studies* **60**, 531–542.
- MELENBERG, B. and A. H. O. VAN SOEST (1994), Parametric and semiparametric estimation of ordered response models with an application to satisfaction with life, mimeo, Tilburg University.

- NELSON, J. (1993), Household equivalence scales: theory versus policy?, *Journal of Labor Economics* 11, 471–493.
- POLLAK, R. A. and T. J. WALES (1979), Welfare comparisons and equivalence scales, *American Economic Review, Papers and Proceedings* 69, 216–221.
- POULIN, S. (1988), An application of analytic techniques to Canadian income satisfaction data, Staff report, Statistics Canada, Labour and Household Surveys Analysis Division, Ottawa.
- ROBINSON, P. (1988), Root-N consistent semi-parametric regression, *Econometrica* 56, 931–954.
- RUUD, P. A. (1984), Tests of specification in econometrics, *Econometric Reviews* 3, 211–242.
- STOCK, J. (1991), Nonparametric policy analysis: an application to estimating hazardous waste cleanup benefits, in: W. A. Barnett et al. (eds.), *Nonparametric and semi-parametric methods in econometrics and statistics*, Cambridge University Press, Cambridge (Mass.).
- VAN PRAAG, B. M. S. (1968), *Individual welfare functions and consumer behavior*, North-Holland, Amsterdam.
- VAN PRAAG, B. M. S. (1991), Ordinal and cardinal utility: an integration of the two dimensions of the welfare concept, *Journal of Econometrics* 50, 69–89.
- VAN PRAAG, B. M. S. and N. L. VAN DER SAR (1988), Household cost functions and equivalence scales, *Journal of Human Resources* 23, 193–210.
- VAUGHAN, D. (1984), Using subjective assessments of income to estimate family equivalence scales: a report on work in progress, *Proceedings of the Social Section of the American Statistical Association*, 496–501.

Received: April 1995. Revised: August 1995.