SPECIFICATION OF HOUSEHOLD ENGEL CURVES BY NONPARAMETRIC REGRESSION

by

Herman J. BERENS
Free University, Department of Econometrics, 1081 BV Amsterdam

and

Hette A. POTTER-BUTER
University of Amsterdam, Department of Economics, 1011 NH Amsterdam

ABSTRACT

This paper demonstrates the usefulness of nonparametric regression analysis for functional specification of household Engel curves.

After a brief review in Section 2 of the literature on demand functions and equivalence scales and the functional specifications used, we first discuss in Section 3 the issue of using income versus total expenditure, the origin and nature of the error terms in the light of utility theory, and the interpretation of empirical demand functions. We shall reach the unorthodox view that household demand functions should be interpreted as conditional expectations relative to prices, household composition and either income or the conditional expectation of total expenditure (rather than total expenditure itself), where the latter conditional expectation is taken relative to income, prices and household composition. These two forms appear to be equivalent. This result also solves the simultaneity problem: the error variance matrix is no longer singular. Moreover, the errors are in general heteroskedastic.

In Section 4 we discuss the model and the data, and in Section 5 we review the nonparametric kernel regression approach.

In Section 6 we derive the functional forms of our household Engel curves from nonparametric regression results, using the 1980 Budget Survey for the Netherlands, in order to avoid model misspecification. Thus the model is derived directly from the data, without restricting its functional form. The nonparametric regression results are then translated to suitable parametric functional specifications, i.e., we choose parametric functional forms in accordance with the nonparametric regression results. These parametric specifications are estimated by least squares, and various parameter restrictions are tested in order to simplify the models. This yields very simple final specifications of the household Engel curves involved, namely linear functions of income and the number of children in two age groups.

Copyright © 1990 by Marcel Dekker, Inc.
1. INTRODUCTION

Household expenditure patterns differ across households according to family size, age composition, educational levels and other household characteristics. In modelling household demand one should therefore not only relate specific expenditures to commodity prices and income or total expenditure, but also to these household characteristics. These models are often used as a basis for welfare comparison between households and the estimation of the cost of children, directly or indirectly via the construction of household equivalence scales.

The various econometric approaches to estimating household demand functions and equivalence scales all have a major drawback in common, namely that the functional form of the demand equations has to be specified in advance, directly or indirectly via the specification of the functional form of the utility function or the cost function. The functional form of the model or the utility function is usually chosen on the basis of tractability rather than on the basis of a priori knowledge of the true functional form. Tractability and reality, however, need not coincide in practice. Since there is almost a continuum of theoretically admissible functional forms, the actually chosen functional form is almost rarely misspecified. This situation is reminiscent of drawing a random variable from a continuous distribution, i.e., the probability that this random variable equals a certain fixed value is equal to zero.

Misspecification of the functional form of household expenditure functions may have serious consequences for the econometric results. In particular, functional misspecification usually leads to inconsistent parameter estimates, and consequently the estimated equivalence scales are inconsistent too. This calls for using methods that are robust for misspecification of functional form. Since the only way to avoid misspecification of functional form is to specify no functional form at all, we shall use nonparametric kernel regression in order to let the functional form of our household Engel curves be determined by the data. Nonparametric kernel regression analysis is a technique which allows consistent estimation of a regression model without specifying in advance its functional form. The only specification that is involved concerns the choice of the dependent variable and the independent variables. The relationship between this dependent variable and the-
independent variables is left free, apart from some mild regularity conditions (such as continuity). On the basis of the nonparametric regression results we shall derive suitable parametric functional specifications and re-estimate and test the models in the traditional way. Our analysis arrives at the conclusion that the data supports the linear Engel model over a wide income range.

2. A SAMPLE FROM THE LITERATURE

The review of the literature below merely aims to give an impression of the variety of functional specifications of household demand functions and Engel curves. It is a sample from the literature rather than a complete overview. For much more complete reviews, see Deaton and Muellbauer (1980) and in particular the excellent reviews by Deaton (1985) and Blundell (1988).

An important concept in household demand theory is the (adult) equivalence scale. Loosely speaking, an adult equivalence scale measures the "needs" of each type of household member in units of the "needs" of a single adult. These scales are in general commodity-specific: it is clear that the need of a baby, relative to the need of an adult, for milk and say, beer, will be quite different. Pollak and Wales (1979) call these scales conditional equivalence scales.

The founding father of household expenditure analysis is Engel (1893). Engel's equivalence scales are based on the proportion of income used for food of Belgium factory workers. The method assumes that the welfare of two households is equal if they spend the same proportion of their income on food. More modern applications assume a common equivalent scale \( r_i(h) \) depending on household composition \( h \). If \( p_i \) is the household expenditure on commodity \( i \), and \( M \) is household income, then \( p_i/(M/r_i(h)) \) is assumed to be a stable function of \( M/r_i(h) \). Following Engel the first applicants of the method used equivalence scales \( r_i(h) \) based on nutritional requirements determined by experts. Although the theory is rather restrictive, as the scale is the same for each commodity, the method has been widely used since the beginning of this century all over the world.

Modern empirical investigations of the expenditure behavior of households are usually based on utility theory, assuming that households act as homogeneous units, with
utility functions depending (directly or indirectly via the cost function or Marshallian demand functions) on household composition. The direct approach specifies the utility function $U$ as a function of a commodity vector $q = (q_1, \ldots, q_N)$ and a vector $h$ of household characteristics: $U = U(q, h)$. The indirect approach is more common and specifies either the cost function $c(p, q, h)$ as a function of a price vector $p = (p_1, p_2, \ldots)$, utility level $u$ and household composition $h$, or the Marshallian demand functions $q(p, M, h)$ as functions of $p$, total expenditure $M$ and household composition $h$. The direct approach and the indirect one via cost functions are of course equivalent, as by the Shephard–Uzawa duality theorem we can recover $U(q, h)$ from $c(p, q, h)$, provided the cost function involved satisfies the requirements for being a proper cost function. The indirect approach via the Marshallian demand functions is also equivalent to the direct approach, provided the demand functions are homogeneous of degree zero in $M$ and $p$, satisfy the adding-up restriction and the corresponding Slutsky matrix is symmetric and negative semi-definite. Cf. Hurwicz and Uzawa (1977).

Given the cost function $c(p, q, h)$ and a reference household of type $k$, one can now answer the question how much more (or less) a household $i$ has to spend in order to maintain the same utility level $u$ as the reference household. For example, if the difference between $i$ and $k$ is just one child, then $c(p, q, h) / c(p, q, h_k)$ may be interpreted as the relative cost of that child. Pollak and Wales (1979) call this ratio an unconditional equivalence scale. This approach assumes that the underlying utility function is a cardinal measure of household welfare, which is disputable. For example, for each household we may transform the utility function $U(q, h)$ into a household specific utility function $U_i(q, h) = F_i(U(q, h))$, with $F_i$ an arbitrary, monotonic increasing function, without affecting the implied demand patterns. See also Fisher (1987) and the discussion in Section 3.

A typical example of the indirect approach via Marshallian demand functions is the model of Pras (1952) and Pras and Houthakker (1953). Pras and Houthakker assume Marshallian demand functions of the form

$$q_i/h(k) = U(p, M/u(h)), i = 1, 2, \ldots, N.$$
where \( \eta_i(h) \) is the commodity-specific equivalent scale of commodity \( i \) and \( \eta(h) \) is the general (income) equivalent scale. The commodity-specific equivalence scales are functions of household composition only. The general or income scale \( \eta(h) \) can be expressed as a function of the specific commodity scales, because of the budget restriction. Since the income scale can be expressed in terms of the specific scale it appears that only the latter need to be estimated. However, Forsyth (1960) discovered that it is impossible to estimate the complete set of specific equivalence scales on the basis of a single budget survey. See Cramer (1969) for a clear exposition of the problem involved. The authors, using semi-, double logarithmic and log-reciprocal Engel curves to pre-war British data, did not seem aware that they would be unable to estimate the specific scales without the implied imposed restriction. The Pras–Houthakker model has also been criticized by Muehlbauer (1984) who argues that if the model is interpreted in terms of utility theory it is consistent with a Leontief utility function only, hence no substitution between commodities is possible.

As Pras (1953) admitted, the approach of Pras (1953) and Pras–Houthakker (1955) is not new. Sylvestrecker and King (1921) were the first to envisage the possibility of incorporating household composition as a variable in Engel curves by weighting with specific equivalence scales for particular commodities.

Another well-known example of the indirect approach is Barten’s (1964) model. Barten assumes Marshallian demand functions of the form

\[
q_i/s(h) = f_i(M_i/p_i(h)) \cdots M_N/q_N(h)));
\]

where \( y_i(h) \) is the specific equivalence scale of commodity \( i \). These demand functions correspond to a utility function of the form

\[
U(q, h) = U(q_i/y_i(h)) \cdots q_N/y_N(h));
\]

This is also the form assumed by Jorgenson and Stienack (1984). A disadvantage of Barten’s model is that it assumes an excessive substitution effect as a result of changes in...
the household composition. Moreover, there are important types of behavior that the model cannot accommodate. In particular, if the reference household (without children) does not consume the good, neither will the household with children, except through the operation of substitution effects. Furthermore, also Bartoš's model suffers from identification problems if only a single budget survey is available for estimation. See Deaton (1986).

Bartoš's model is equivalent to a cost function of the form \( c(p^*, u) \) with \( p^* = p \cdot \eta(h) \). Gorman's (1976) modification amounts to adding a fixed cost term \( p \cdot c_0(h) \) to \( c(p^*, u) \), with \( c_0(h) \) a vector function of \( h \): \( c(p, u, h) = p \cdot c_0(h) + c(p^*, u) \), in order to allow for more flexible substitution responses. This is an example of how to transform an initial cost function \( c(u, p) \) into a cost function depending also on demographic variables. Lewbel (1983) gives general sufficient conditions for incorporating demographic effects into an initial legitimate cost function without demographic variables such that the transformed cost function is theory consistent.

An important and striking result regarding the functional form of demand functions is obtained by Gorman (1981). Gorman considers demand functions of the general form

\[
pq_i/M = \sum_{j=1}^{k} A_{ij}(p)q_j(M), i = 1, \ldots, k
\]

He shows that if such equations are to be theory consistent, the rank of the matrix formed from the coefficients \( A_{ij}(p) \) can be no larger than 3 and the functions \( q_j \) must take specific restricted forms, such as \( q_j(\cdot) = (\cdot)^j \) or \( q_j(\cdot) = \exp((j-1)(\cdot)) \). The latter case yields the quadratic expenditure model

\[
pq_i = \alpha(p) + \beta(p)M + \gamma(p)M^2
\]

See Howe, Pollak and Wales (1979) and Van Dael and Merkies (1989). This model is, of course, a further generalisation of Stone's (1954) famous linear expenditure system (LES).

By letting \( A_{ij}(\cdot) \) depend on household composition \( h \) as well, we can generalize Gorman's approach to household demand functions. See, e.g., Pollak and Wales (1978, 1979a, 1981).
Engel curves focus on the relation between specific expenditure categories $p_i$ and income or total expenditure $M$. Household Engel curves take also the household composition $h$ into account. They follow from the Marshallian demand functions by treating the price vector $p$ as fixed. This is appropriate if one uses a budget survey of a single year, like the survey we shall use. Estimation of a complete demand system and equivalence scales requires panel data with a sufficient number of waves to identify the price effects as well.

A wide selection of functional forms for Engel curves has been explored in the literature. Quoting Deaton and Muellbauer (1980, p.19): "More complex forms, such as the cumulative distribution function of the lognormal distribution, have been suggested that combine many of the desirable features of the simpler forms. None of these forms is fully consistent with adding-up, and although this may be less serious a problem than in time series analysis, the theoretical plausibility of these models is not enhanced by their failure to meet this requirement."

Blundell and Ray (1984) have touted the linear Engel curve specification against a class of nonlinear Engel curves. Their test rejects the linear Engel model. This is at odds with our finding for the Netherlands.

A simple form of Engel curves which is consistent with adding-up was first estimated by Working (1943), and later by Leser (1963), among others. This approach relates budget shares $p_i/M$ linearly to the logarithm of total expenditure $M$: $p_i/M = a_i + \beta \ln(M)$. Clearly, adding-up requires that $X_0 = 1$ and $\sum_i = 0$. The Working-Leser model is, of course, a special form of Gorman's model. The Working-Leser model corresponds to a cost function of the form

$$\ln[c(a,p)] = a - \ln[b(p)] + (1-a) - \ln[a(p)],$$

where $c(p)$ and $b(p)$ are homogeneous of degree one, representing the cost functions of the very poor ($a = 0$) and the very rich ($a = 1$), respectively. A full system of demand functions within the Working-Leser class can be generated by suitable choice of the functions $b(p)$ and $a(p)$, such as the "almost ideal demand system" (AIDS) of Deaton and
Muehlbauer. Extensions of these models to households can be obtained by specifying the functions \( a() \) and \( b() \) as functions of \( p \) and household composition \( h \). A simple extension of the Working-Leeb Engel curve that incorporates demographic effects is chosen by Deaton and Muehlbauer (1986):

\[
p_{j}\ln(M) = \alpha_{j} + \beta_{j} \ln(M/n) + \gamma_{j} h_{j},
\]

where \( n_{j} \) is the number of persons in category \( j \) (\( j = 1, \ldots, J \)) and \( n \) is the total number of persons in the household. For many third world surveys Deaton and Muehlbauer (1986) found that the \( \ln(M/n) \) term provides a high degree of explained variation and that the \( \gamma \) parameters are rather small.


We recall that in empirical household demand analysis the Marshallian demand functions are directly or indirectly derived from a utility maximisation problem:

\[
\max_{q} U(q/h), \text{ subject to } q \cdot p = M,
\]

where \( M \) is total expenditure, \( p = (p_{1}, \ldots, p_{J}) \) is a price vector, \( q = (q_{1}, \ldots, q_{J}) \) is a commodity vector and \( h \) is a vector that measures household composition. Moreover, the functional form of the utility function \( U \) is specified in advance as a known function of \( q, h \) and some parameter. This functional specification of \( U \) then determines the functional specification of the Marshallian household demand functions

\[
q_{j} = g(M, p, h), j = 1, \ldots, k.
\]

However, since these demand functions never fit the data perfectly, an error term is added:
(3.3) \[ q_{it} = g_i(M, p, h) + v_{it}, i = 1, \ldots, k, \quad v = (v_1, \ldots, v_k)^T, \quad E v = 0. \]

Since by definition \( M = p'q \), we have \( P[v'v = 0] = 1 \). The latter condition implies that \( v \) depends on \( p \), hence we cannot make the usual regression assumption that the \( v_i \) are independent of \( p, M \) and \( b \). Moreover, \( E vv' \) is a singular matrix. One could, of course, specify \( e_i = v_i/p_i \), by which the model becomes

(3.4) \[ pq_i = p g_i(M, p, h) + e_i, i = 1, \ldots, k, \quad e = (e_1, \ldots, e_k)^T, \quad E e = 0. \]

The adding-up condition then implies \( P[e_i = 0] = 1 \), so that also now the variance matrix \( E e e' \) is singular. Most researchers now assume that the errors \( e_i \) are independent of \( p, M \) and \( b \), homoskedastic or even normally distributed. One may then proceed by deleting one equation and estimating the remaining \( k-1 \) equations by say SURE.

Two problems now arise. First: How is \( M \) determined? Second: How do the errors \( e_i \) or \( v_i \) fit in utility theory? The solution to the first problem is well known, namely the life cycle consumption model of Modigliani and Brumberg (1955), in particular the versions of Hall (1978) and Winder and Palm (1980) of this model are useful for our purpose. See also Deaton (1986), Blundell (1988) and the references in Winder and Palm (1990). As we will show, the answer to the second problem also follows from the life cycle consumption model.

Following Winder and Palm (1980) we start with assuming that a consumer maximizes an intertemporal additive indirect utility function

\[ U_T(\omega(t), u(c_t | \theta)) \]

subject to the life-time budget constraint

\[ U_T(1+c(T)|c_1) = (1+c(T)\mu + \sum_{t=1}^T (1+c(t))^{-2}E[p_t | F_0]. \]
where

\[ u(c(t)) \] is an indirect utility function corresponding to the utility function \( U(q; \theta) \), with \( \theta \) a parameter vector.

\[ c_t \] is real consumption in period \( t \);

\[ y_t \] is real net income in period \( t \);

\[ w_t \] is the endowment of real wealth at the beginning of period \( t=0 \);

\[ r \] is the real interest rate which is assumed to be constant over life-time and equal for all consumers;

\[ R(\theta) \] is a time preference weight depending on the parameter vector \( \theta \);

\[ T \] is life-time and

\[ F_0 \] is an information set containing \( y_0 \) and \( w_0 \), together with other variables to be specified later. (\( F_0 \) should be interpreted as the Borel field generated by these variables. Cf. Chung 1974).

Thus, the consumer under review is rational in the sense that he estimates his unknown future real income by a conditional expectation.

Assuming that the consumer acts on the basis of constant prices \( p \) and that \( p \) belongs to the information set \( F_0 \), we can restate the model as

\[
\max \sum_{t=0}^{\infty} R(\theta) U(q_t; \theta) \]

subject to

\[
\sum_{t=0}^{\infty} (1+r)^{-t} p_t q_t = (1+r)p_0 w_0 + \sum_{t=0}^{\infty} (1+r)^{-t} q_t R(y_t; F_0),
\]

with \( r(p) \) a price index which is homogeneous of degree one.

If the economic agent involved is a household rather than a single person it is reasonable to assume that \( U(q_t; \theta) \) and \( R(\theta) \) depends on household composition \( h_t \) in period \( t \). Thus the utility function becomes

\[
\sum_{t=0}^{\infty} R(h_t, \theta) U(q_t; h_t, \theta).
\]
However, for a household life-time T is no longer well defined. Since the components of \( h_0 \) are the number of persons of various categories in the household, one may identify the situation where the household ceased to exist with a particular value of \( h_0 \), say \( h_0 = 0 \). Then the model can be restated as:

\[
\max \sum_{l} u_l \{ d_l(h_0, \theta) \} \mathbb{E}[y_l | h_0, \theta] \\
\text{subject to} \\
\sum_{q} \{ d_q \} \{ 1 + r \}^{p_q} q_l = (1 + r) \{ p \} w_q + \sum_{h} \{ d(h_0) \} \{ 1 + r \} \{ (p) \} \mathbb{E}[y_l | F_0],
\]

with \( l(\cdot) \) the indicator function. However, \( h_0 \) is not perfectly predictable. Therefore one should take conditional expectations. Assuming that \( h_0 \) is a discrete random variable taking values in a finite set \( H \) and that \( h_0 \) belongs to \( F_0 \), we have

(3.3a) \[
\max \sum_{l} u_l \{ q_l | h_0, \theta \} \\
\text{subject to} \\
\sum_{q} \{ d_q \} \{ 1 - P_l(0 | h_0) \} \{ 1 + r \}^{p_q} q_l = (1 + r) \{ p \} w_q + \sum_{h} \{ d(h_0) \} \{ 1 + r \} \{ (p) \} \mathbb{E}[y_l | F_0],
\]

where \( P_l(0 | h_0) \) is the conditional probability of \( h_0 \) given \( h_0 \) and

\[
E(q_l | F_0) = \sum_h h E[l(h) | h] \theta / U(q_l | h_0, \theta) P_l(h | h_0).
\]

Note that \( U(q_l | h_0, \theta) = U(q_l | h_0, \theta) \).

In the course of the argument we have assumed that \( F_0 \) contains \( y_b, h_0, w_0 \) and \( p \). Since conditional expectations are Borel measurable functions of the conditioning variables [cf. Chung (1974, Chap. 9)], we can write

\[
E[y_l | F_0] = \eta_l(y_b, h_0, w_0, p).
\]
The functions \( \gamma_i \) are homogeneous of degree zero in \( p_i \), for \( \gamma_i \) is real income and thus invariant with respect to proportional shifts in \( p_i \).

The solution of the problem (3.5) can now be written as

\[
q_{it} = \mu_{it}(y_{i}, b_{i}, w_{i}, p_{i}), \quad i=1, \ldots, k, \quad t=0,1, \ldots,
\]

and these functions \( \mu_{it} \) are homogeneous of degree zero in \( p_i \). We are particularly interested in the solution for \( t=0 \), as the quantity \( M \) in (3.1) now equals

\[
M = \sum_{i} p_{i} \mu_{i}(y_{i}, b_{i}, w_{i}, p_{i}).
\]

Obviously, for \( t=0 \) the system (3.6) is also the solution of problem (3.1) with \( U(q_{i}|b) \) replaced by \( U(q_{i}|b_{i}) \). Moreover, \( w_{i} \) enters the demand functions for \( t=0 \) via \( M \) only.

Denoting

\[
Y = \pi(p)y_{1} = \text{nominal net current income},
\]

and dropping the subscript \( i \) we now obtain household demand functions for \( t=0 \) of the type

\[
q_{i} = \phi(Y, p, b, w, \theta) = g_{i}(M, p, b, \theta), \quad i=1,2, \ldots, k,
\]

with total expenditure

\[
M = \sum_{i} p_{i} \phi_{i}(Y, p, b, w, \theta).
\]

Note that the demand functions \( \phi_{i} \) in (3.7) are homogeneous of degree zero in \( Y \) and \( p_i \), and that (3.8) is homogeneous of degree one in \( Y \) and \( p_i \).

Arriving at this stage traditional empirical household demand analysis will now be concerned with the estimation of the unknown parameter vector \( \theta \). However, this implicitly assumes that \( \theta \) is the same for all households, i.e., households facing the same
values of $Y$, $p$, $h$ and $w$ will buy the same commodity bundle $q$. It is undeniable that this assumption conflicts with reality. Thus, the parameter vector $\theta$ actually varies from household to household. The same applies to the endowment $\omega$ of real wealth, which is often not observed. Consequently $\omega$ and $w$ actually play the role of disturbance terms. Denoting by $F(w, \delta | Y, p, h)$ the conditional distribution of $(w, \delta)$ relative to $(Y, p, h)$ we can now write:

\begin{align}
(3.9a) \quad q_i &= f(Y, p, h) + \nu_i, \quad i = 1, \ldots, k. \\
\text{where} \quad f(Y, p, h) &= \int f_i(Y, p, h, w, \delta) dF(w, \delta | Y, p, h) \\
(3.9b) \quad \nu_i \text{ is an error term satisfying} \quad E[\nu_i | Y, p, h] = 0 \text{ with probability 1.} 
\end{align}

Note, however, that the conditional variance matrix of $\nu = (\nu_1, \ldots, \nu_k)'$ is not necessarily constant: $E[\nu' \nu | Y, p, h]$ will in general be a matrix function of the conditioning variables involved.

Similarly, we can write

\begin{align}
(3.10a) \quad M &= m(Y, p, \omega) + p' \nu \\
\text{where} \quad m(Y, p, \omega) &= \int m_i(Y, p, \omega, w, \delta) dF(w, \delta | Y, p, h). \\
(3.10b) \quad \nu &= m(Y, p, h).
\end{align}

Next, denote

\begin{align}
(3.11) \quad X &= m(Y, p, h).
\end{align}
From the life cycle consumption model it follows that under mild conditions \( M \) is a strictly monotonic increasing function of \( Y \), given \( p, h, w \) and \( \theta \). Thus for \( \lambda > 1 \),

\[
\Omega_{\lambda, y, p, h, w, \theta} > \Omega_{\lambda, y, p, h, w, \theta}.
\]

Moreover, if \( \lambda \) is fixed stochastic \( \lambda \), \( (Y, p, h) \) and \( (Y, p, h) \) generate the same Borel field, hence conditioning on \( (Y, p, h) \) is equivalent to conditioning on \( (Y, p, h) \). Cf. Chung (1974). Consequently,

\[
F(w, \theta) = F(w, \theta) \quad \text{with probability 1}.
\]

Combining (3.12) and (3.13) we now see that also \( m(Y, p, h) \) is strictly monotonic in \( Y \), in the sense that for stochastic \( \lambda > 1 \), \( m(\lambda Y, p, h) > m(Y, p, h) \) with probability 1, hence conditioning on \( (Y, p, h) \) is equivalent to conditioning on \( (X, p, h) \), as \( (Y, p, h) \) and \( (X, p, h) \) generate the same Borel field. Therefore we may write the demand system (3.9) also as

\[
q_i = g_i(X, p, h) + v_i, i = 1, \ldots, k,
\]

or

\[
pq_i = y g_i(X, p, h) + c_i, \quad i = 1, \ldots, k, \quad c_i = v_i / p_i.
\]

Note that \( (Y, p, h) \) and \( (M, p, h) \) do not generate the same Borel field, due to the fact that \( M \) depends on \( w \) and \( \theta \). Therefore the equivalence of (3.14) and (3.9a) breaks down if we condition on \( (M, p, h) \) instead of \( (X, p, h) \). Of course, nobody can forbid us to condition on \( (M, p, h) \) instead of \( (Y, p, h) \) or \( (X, p, h) \), but then the error of models (3.3) and (3.4) prevents us from interpreting actual demand \( q \) as the outcome of a deterministic utility maximization problem, and the same applies to the theoretical demand functions \( q_i = g_i(M, p, h) \), for adding-up does not hold for the \( q_i \).

Comparing models (3.3) and (3.4) with (3.14) and (3.15) we see two fundamental differences. First the role of \( M \) in (3.3) and (3.4) has been taken over by the conditional
exposition $X = E[M|Y,p,h]$. Second, the variance matrices $E v\pi$ and $E x\pi$ of the errors in models (3.14) and (3.15) are no longer necessarily singular. Moreover, from the way we have derived these models it is clear that in general $E[v\pi|Y,p,h]$ and $E[x\pi|Y,p,h]$ are non-constant matrix functions of $Y$, $p$, and $h$, c.q. of $X$, $p$, and $h$. Thus the errors $e_i$ and $v_i$ are in general heteroscedastic.

Finally, the question arises whether the functions $g_i(X,p,h)$ can be interpreted as solutions of a utility maximisation problem, i.e., does there exist a deterministic utility function $U^t(q|h)$ such that the functions $g_i(X,p,h)$ are solutions of the problem max $U^t(q|h)$ subject to $p^t=q=1$? Obviously, adding-up does apply by construction. But what about homogeneity? The conditional distribution $F(w|Y,p,h)$ is homogeneous of degree zero in $Y$ and $p$. This follows from the fact that for given non-stochastic $\lambda > 0$ the Borel fields generated by $(AY,Ap,h)$ and $(Y,p,h)$, respectively, are the same, hence similarly to (3.13), $F(w,d|Y,p,h) = F(w,0|Y,p,h)$ with probability 1. Since $M$ and the functions $g_i(Y,p,h,w,\theta)$ are homogeneous in $Y$ and $p$ of degree one and zero, respectively, it follows that $X = m(Y,p,h)$ is homogeneous of degree one in $Y$ and $p$ and the functions $f_i(Y,p,h)$ are homogeneous of degree zero in $Y$ and $p$. Since conditioning on $(Y,p,h)$ is equivalent to conditioning on $(X,p,h)$, it follows now that the functions $g_i(X,p,h)$ are homogeneous of degree zero in $X$ and $p$. With this result at hand it suffices to verify the symmetry and negative semi-definiteness of the Slutsky matrix. See Hurwicz and Uzawa (1971). However, without explicit functional forms for the $g_i(X,p,h)$, verifying the properties of the Slutsky matrix does not seem feasible.

The conclusions from the above argument may now be summarized as follows:

1) The disturbance terms in empirical demand systems originate from (randomly) varying preferences and unobservable variables like real wealth.

2) Demand systems that relate specific demand to total expenditure $M$, prices $p$ and household composition $h$ are either not identifiable because the households have different preferences, or should fit the data without error.

3) Consequently, interpersonal and interhousehold welfare comparison on the basis of models that do not fit the data very well does not make sense.
4) The life cycle consumption hypothesis leads to demand systems that relate specific
demand to net current income Y, prices p and household composition h, plus a
heteroscedastic error term. The demand functions involved are conditional expectation
functions.

5) The conditional expectation functions involved can be rewritten as functions of X, p
and h instead of Y, p and h, where X is the conditional expectation of total expenditure M,
relative to Y, p and h. The transformed conditional expectation functions are
homogeneous of degree zero in X and p and satisfy the adding-up condition.

4. MODEL AND DATA

The household Engel curves we consider relate expenditures of household j on a
certain group of commodities to net income (including child-allowances) of household j,
the number of children in the age group 0-15 and the number of children in the age group
16 or over in household j. The latter only concerns children living with their parents and
having no income themselves. Since this study merely aims to be a pilot study of the
applicability of nonparametric regression analysis in the empirical area under review, we
keep the analysis here as simple as possible by distinguishing only two expenditure
categories, namely

\[ y_{ij} = \text{expenditures of food, clothing and footwear}, \]
\[ y_{ij} = \text{other expenditures} \]

of household j. For the very same reason we only distinguish two age groups. The
explanatory variables are now:

\[ x_{ij} = \text{net income}, \]
\[ x_{ij} = \text{number of children in the age group 0-15}, \]
\[ x_{ij} = \text{number of children in the age group 16 or over} \]
of household j. The household Engel models involved are:
where the response functions or regression functions \( g_1 \) and \( g_2 \) are completely unknown, apart from the condition that \( g_1 \) and \( g_2 \) are continuously differentiable in \( x_4 \). By construction the disturbance terms \( u_1 \) and \( u_2 \) satisfy the usual condition that their conditional expectations relative to the regressors \( x_1, x_2 \), and \( x_3 \) equal zero with probability 1:

\[
E(u_1|x_1, x_2, x_3) = 0 \quad \text{and} \quad E(u_2|x_1, x_2, x_3) = 0 \quad \text{with prob. 1.}
\]

Note that in view of the argument in Section 2 these errors are likely heteroskedastic.

The existence of \( g_1 \) and \( g_2 \) is guaranteed by the conditions \( E(y|u) < 1 \) and \( E(y_2|u) < 1 \). Cf. Chang (1974, Theorem 9.1.1). Moreover, the functions \( g_1 \) and \( g_2 \) are unique, given the i.i.d. data generating process, in the sense that if there exist other functions \( f_1 \) and \( f_2 \), respectively, with the above properties, then \( P(g_j(x_1, x_2, x_3) = f_j(x_1, x_2, x_3)) = 1, i=1,2 \). However, if we would extend the list of variables the conditional expectations functions will in general change, and the same applies if we would transform the dependent variables \( y_2 \).

Thus, the uniqueness of \( g_1 \) and \( g_2 \) is conditional on the choice of the dependent variables \( y_2 \) and the regressors. For example, if we would take \( \ln(y) \) and \( \ln(y_2) \) instead of \( y_1 \) and \( y_2 \), respectively, we get different models with different interpretations. On the other hand, taking \( \ln(x_4) \) instead of \( x_4 \) does not matter, as conditioning on \( (x_1, x_2, x_3) \) is equivalent to conditioning on \( \ln(x_4) \).

\[
E(y_2|\ln(x_4), x_2, x_3) = E(y_2|x_1, x_2, x_3) = g(x_1, x_2, x_3).
\]

Of course, the uniqueness of \( g_1 \) and \( g_2 \) only applies to cross-section data. The household Engel curves will likely change over time due to changes in preferences and prices.

We recall that no assumptions about the functional form of \( g_1 \) and \( g_2 \) will be made. We only assume that the variable \( x_{1i} \), net income, is continuously distributable and that \( g(x_1, x_2, x_3) \) and \( g(x_1, x_2, x_3) \) are for each pair \( (x_2, x_3) \) continuously differentiable in \( x_1 \).
The data set we work with is the 1980 Budget Survey held by the Dutch Central Bureau of Statistics. This survey consists of an independent sample of 2859 households. For technical reasons we have split this sample in two subsamples of sizes 2000 and 859, respectively. The smaller subsample has been used for experiments with the nonparametric regression method, in order to improve the fit. Cf. Section 5.4. The larger subsample has been used for the actual nonparametric estimation of our Engel functions.

A typical feature of the budget survey involved is that total expenditures may exceed net income, especially in the low income range. This is due to the fact that expenditures on durables are completely attributed to the year of purchase. Thus, if a household buys, say, new furniture, in a certain year, the total amount of the purchase involved is considered as an expenditure in that year, even if the purchase has been financed by a loan. The same applies to clothing and footwear: although a suit or a pair of shoes may last longer than a year, the total amount of the purchase is considered as expenditures in the year of the purchase. However, this is not a serious problem. The life cycle model can be modified to include durables so that the general argument in Section 1 goes through. So the interpretation of the model is the same with or without durables. Moreover, the nonparametric approach is in principle capable of picking up the consequence of the presence of durables for the functional form of our household Engel curves. Furthermore, our aggregation level is such that zero expenditures do not occur. Cf. Keen (1986).

Since the 1980 Budget Survey is a representative survey, it also contains households with only one parent and households of elderly. These households have been excluded from our analyses (after splitting the sample in two subsamples). However, the remaining data subsets of sizes 1130 and 552, respectively, are then no longer random samples, a situation not accounted for in the theory of nonparametric estimation. As will be shown in Section 5, a simple modification of the nonparametric regression approach will correct for sample selection.

Finally we note that the further subsample of size 1130 contains five households with expenditures on food, clothing and footwear exceeding net income, 86 households with other expenditures exceeding net income and 424 households with total expenditure
exceeding net income. For the further subsample of size 522 these numbers are 1, 48 and 226, respectively. This is mainly due to the typical way expenditures are measured in the budget survey under review, although we do not exclude that also occasional measurement errors in net income may contribute to this phenomenon (despite the assurance of CBS that in the survey under review income is accurately measured). However, measurement errors in income cannot be taken into account in a nonparametric framework. So, if there are any, we shall ignore them.

5. THE KERNEL REGRESSION FUNCTION ESTIMATOR

5.1 Introduction

In this section we review the asymptotic properties of the Nadaraya-Watson type kernel estimator of an unknown (multivariate) regression function. Since the pioneering papers of Nadaraya (1964) and Watson (1964) on kernel regression function estimators there is now a growing extent of literature on the problem of nonparametric estimation of unknown regression functions. See Clossm (1981, 1985), Bierens (1987) and Hardle (1989) for bibliographies. Most of the literature on nonparametric regression function estimation deals with the kernel method and its variants.

Consider an i.i.d. sample \( \{(y_i, x_i)_\ldots(x_n,\ldots, x_0)\} \). In this data set the \( y_i \)'s are the dependent variables and the \( x_i \)'s are \( k \)-component vectors of regressors. If \( E|y_i| < \infty \) then the conditional expectation of \( y_i \) relative to \( x_i \) exists and takes the form

\[
E(y_i|x_i) = g(x_i),
\]

with \( g(\cdot) \) a Borel measurable real function on \( \mathbb{R}^k \). Cf. Chang (1974). Denoting

\[
u_i = y_i - g(x_i),
\]

we then get the regression model

\[
x_i = g(x_i) + u_i,
\]
where by construction the error term $u_j$ satisfies the usual condition that its conditional expectation relative to the vector of regressors equals zero with probability 1, i.e.,

$$(5.4) \quad P(E(u_j|x_j) = 0) = 1.$$ 

The model (5.3) is therefore purely tautological, that is, its set up is merely a matter of definition. Now our aim is to estimate the regression function $g(x)$ without making explicit assumptions about its functional form. Here we use the kernel regression approach. There are various other techniques that can accomplish the same task, such as nearest neighbor regression [Stone (1977)], series expansion approaches like Gallant’s Fourier transform approach [cf. Gallant (1981, 1985, 1986) and El Badavi, Gallant and Souza (1983)] and Barnett’s (1983) Laurent series approach, spline functions [see, e.g., Wabba (1978) and Wahba and Wold (1975) for the statistical theory involved and Blandell (1980) for an application], and projection pursuit [cf. Friedman and Stuetze (1981) and Huber (1985)].

The reason for working with the kernel approach is that it can easily handle mixed continuous-discrete regressors and the asymptotic theory involved is well-developed. Cf. Bierens (1982, 1987). See also Hardle (1990) for a review of various nonparametric regression techniques and Hardle, Hildebrand and Jeroson (1988) for another application of kernel estimation to demand theory.

Although the distribution of our vector of regressors is mixed continuous-discrete, we first discuss the pure continuous case as an introduction to the general case. Thus, for the moment we shall assume that $x_j$ is continuously distributed with density $f(x)$. In Section 5.3, we consider the mixed continuous-discrete case.

The Nadaraya-Watson kernel regression function estimator of $g(x)$, named after Nadaraya (1964) and Watson (1964), is a random function of the form

$$(5.5) \quad \hat{g}(x) = \frac{1}{n} \sum_{j=1}^{n} K\left(\frac{x-x_j}{hn}\right) \frac{y_j}{h} = \frac{\hat{g}(x)f(x)}{f(x)}.$$ 

say, with
the kernel estimator of the marginal density \( f(x) \) of \( x_i \). The latter estimate has been proposed by Rosenblatt (1956). Important contributions to the asymptotic theory of this class of estimators have been made by Parzen (1962) for the univariate case and Cacoullos (1966) for the multivariate case. See Fryer (1977) and Tapia and Thompson (1978) for reviews.

The function \( K(.) \) is an a priori chosen real function on \( R^d \), called the kernel, satisfying

\[
\int |K(x)| dx < \infty, \quad \int K(x) dx = 1, \tag{3.7}
\]

and \( \gamma_n \) is an a priori chosen sequence of positive numbers, called window width parameters, satisfying

\[
\lim_{n \to \infty} \gamma_n = 0, \quad \lim_{n \to \infty} n \gamma_n^d = a. \tag{3.8}
\]

Note that the kernel regression function estimator is a weighted mean of the dependent variables \( y_i \), where the weights sum up to 1. In particular, if the kernel is chosen to be a unimodal density function with zero mode, for instance let the kernel be the density of the \( k \)-variate standard normal distribution, then the closer \( x \) is to \( x_i \), the more weight is put on \( y_i \).

Under conditions (3.7) and (3.8) the estimator \( \hat{f}(x) \) is pointwise consistent in every continuity point of \( f(x) \), provided

\[
\sup_{x} E(x) < a. \tag{3.9}
\]

The proof of this proposition is simple but instructive. First, the asymptotic unbiasedness follows from
\[ E \tilde{f}(s) = \int \gamma_n K((x-s)/\gamma_n)K(s)dx = \int E(x-\gamma_n x)K(s)dx \]
\[ = E \tilde{f}(x)/K(s)dx = 0 \]
by bounded convergence. Second, the variance vanishes at order \(O(1/\gamma_n^2))\), as
\[
\begin{align*}
\text{var}(f(x)) &= n^{-2/\varepsilon}(1/n^2)\text{var}(K((x-x_0)/\gamma_n)) \\
&= E \gamma_n^2 K((x-x_0)/\gamma_n)^3 - 3\gamma_n^2 E \gamma_n K((x-x_0)/\gamma_n) \\
&= E g(x)K((x-x_0)/\gamma_n^3)dx - 3\gamma_n^2 E g(x)K(x)dx \\
&= E g(x)K((x-x_0)/\gamma_n)K(s)dx - E g(x)K(x)dx
\end{align*}
\]
by bounded convergence. This completes the pointwise consistency proof.

Similarity to (5.10) and (5.11) it follows now that
\[
\text{var}(g(x)) = E \gamma_n^4 K((x-x_0)/\gamma_n)^4 - 4\gamma_n^4 E \gamma_n K((x-x_0)/\gamma_n) \\
= E (\gamma_n^2 K((x-x_0)/\gamma_n)^3 - 3\gamma_n^2 E \gamma_n K((x-x_0)/\gamma_n)) \\
= E (\gamma_n^2 K((x-x_0)/\gamma_n)^3 - 3\gamma_n^2 E \gamma_n K(x)) \\
= E g(x)^3 K((x-x_0)/\gamma_n)K(s)dx - 3\gamma_n^2 E g(x)^2 K(s)dx \\
= E g(x)^3 K(s)dx - 3\gamma_n^2 E g(x)^2 K(s)dx
\]
by bounded convergence, where
\[
\sigma^2(x) = E(|y - x|^2) \text{ for } f(x) > 0,
\]
provided \(x\) is a continuity point of \(g(x), f(x)\) and \(\sigma^2(x)\) and
\[
\sup g(x)|f(x)| < \infty, \sup \sigma^2(x)|f(x)| \infty.
\]
Now it is easy to verify from (5.10) through (5.13) that

\[(\mathbf{5.16}) \quad \lim_{n \to \infty} \hat{f}(x) = f(x), \quad \lim_{n \to \infty} g(x) \hat{f}(x) = g(x)f(x)\]

and hence

\[(\mathbf{5.17}) \quad \lim_{n \to \infty} g(x) \hat{f}(x) = g(x)\]

in every continuity point x of f(x) and g(x)f(x) for which f(x) > 0.

The consistency of the kernel regression function estimator is not limited to the case that \(x_i\) is continuously distributed, as is shown by Devroye (1978), Devroye and Wagner (1980) and Bierens (1983). We shall consider the mixed continuous-discrete case later on, in section 5.3.

### 1.7 Asymptotic normality in the continuous case

The kernel regression estimation approach distinguishes itself from other nonparametric regression methods in that asymptotic distribution theory is fairly well established. In particular, the asymptotic normality of the kernel regression function estimator for the case of continuously distributed regressors has been proved by Schuster (1972) for the univariate case (k=1). Bierens (1987) gave a simple proof for the general case k ≥ 1. Bierens' argument amounts to the following: Let for \(p > 0\),

\[(\mathbf{5.18}) \quad \sigma^2(x) = E[|u_j|^p | x_j = x],\]

provided \(E|u_j|^p < \infty\) and \(f(x) > 0\), and assume:

\textbf{ASSUMPTION 5.1.} There exists a \(\delta > 0\) such that \(\sigma^2(x) f(x)\) is uniformly bounded. The functions \(g(x) f(x)\) and \(\sigma^2(x) f(x)\) are continuous and uniformly bounded. The functions \(f(x)\) and \(g(x) f(x)\) and their first and second partial derivatives are continuous and uniformly bounded.
Now observe that

\[
\begin{align*}
(\dot{g}(x)-g(x))f(x) &= \int (1/\gamma_0)E\left[\ell \left( \frac{x-x_0}{\gamma_0} \right) \right] K(x-x_0) \dot{\gamma}_0 \\
&+ \int (1/\gamma_0)E\left[\ell \left( \frac{x-x_0}{\gamma_0} \right) \right] K(x-x_0) \gamma_0 \\
- B(g(x_0)-g(x))K\left(\frac{x-x_0}{\gamma_0} \gamma_0 \right) \\
&+ \int (1/\gamma_0)E\left[\ell \left( \frac{x-x_0}{\gamma_0} \right) \right] K(x-x_0) \gamma_0 \\
= & \dot{\gamma}(x) + \dot{\gamma}(x) + \dot{\gamma}(x),
\end{align*}
\]

say. It follows from Liapunov's central limit theorem that

\[
\sqrt{n}/\gamma_0 \hat{\gamma}(x) \to N(0, c_0^2 f(x) \int K(x)^2 \, dx) \quad \text{in distr.},
\]

where for \( f(x) > 0 \)

\[
\hat{\sigma}_n^2(x) = \text{E}(\dot{u}_n^2 | x_j = x)
\]

is the conditional variance of \( u_n \). The asymptotic variance in (5.20) can be derived in the same way as in (5.11). Also in the same way as in (5.11) it can be shown that

\[
\lim_{n \to \infty} \text{E} \left( \sqrt{n} \gamma_0 \hat{\gamma}(x) \right)^2 = 0.
\]

Finally, observe that similarly to (5.10)

\[
\begin{align*}
\dot{\gamma}(x) &= \int (g(x)-\gamma_0 x) f(x) \, dx \\
&= \int (g(x)-\gamma_0 x) f(x) \, dx - g(x) f(x) \, dx \\
&= \int (g(x)-\gamma_0 x) f(x) \, dx \\
&+ \frac{\gamma_0}{2} \int x ((\partial/\partial x) + (\partial/\partial x) \gamma_0) f(x) \, dx \\
&+ \gamma_0 g(x) \int x ((\partial/\partial x) + (\partial/\partial x) \gamma_0) f(x) \, dx \\
&- \frac{\gamma_0}{2} \int g(x) f(x) \, dx \\
&- \frac{\gamma_0}{2} \int g(x) f(x) \, dx,
\end{align*}
\]

where \( 0 \leq \gamma_0(x,a) \leq 1 \). The last equality in (5.23) follows from Taylor's theorem. Thus if
we choose $K$ such that

\[(5.24) \quad \int xK(x)dx = 0, \int x^2K(x)dx = \Omega \text{ is finite}, \]

then the first and third terms at the right hand side of (5.23) vanish, while by bounded convergence the second and fourth terms, divided by $\gamma_n$, converge. Thus,

\[(5.25) \quad \lim_{n \to \infty} \gamma_n^2 g(x) = b(x),\]

where

\[(5.26) \quad b(x) = \frac{\exp \left\{ \int (h(x)g(x)) \right\}}{\exp \left\{ \int (h(x)g(x)) \right\}} \cdot \exp \left\{ \int \left( \frac{\partial}{\partial x} h(x) \right) \right\} \cdot \exp \left\{ \int \left( \frac{\partial}{\partial x} g(x) \right) \right\}.\]

From these results it follows:

**Theorem 5.1.** Let assumptions 1.1 and condition (5.21) hold and let $E(x) > 0$. If

\[(5.27) \quad \lim_{\mu \to \infty} \frac{1}{\mu} \| \varphi \| = \rho \text{ with } 0 < \rho < \mu\]

then

\[(5.28) \quad \varphi^{(n)}(x) = g(x) - \frac{\exp \left\{ \int (h(x)g(x)) \right\} - \exp \left\{ \int (h(x)g(x)) \right\}}{\exp \left\{ \int (h(x)g(x)) \right\}} \cdot \exp \left\{ \int K(x)dx \right\}\]

in distribution. If

\[(5.29) \quad \lim_{\mu \to \infty} \frac{1}{\mu} \| \varphi \| = 0,\]

then

\[(5.30) \quad \lim_{\mu \to \infty} \frac{1}{\mu} \| \varphi \| = b(x)/E(x)\]
Note that the latter result may be considered as convergence in distribution to a degenerated limiting distribution.

At first sight it looks attractive to choose the window width \( \gamma_n \) such that \( \mu = 0 \), as then the asymptotic bias vanishes. However, in that case the asymptotic rate of convergence in distribution is lower than in the case \( \mu > 0 \), as (5.27) implies

\[
\sqrt{2(\log n)^{2/(k+4)}} \sim \mu^{1/(k+4)} \text{ as } n \to \infty.
\]

Thus the window width \( \gamma_n \) which gives the maximum rate of convergence in distribution is

\[
\gamma_n = c\mu^{-1/(k+4)},
\]

where \( c > 0 \) is a constant. Since \( \mu = \sigma^{(k+4)/2} \), we have the following corollary.

**THEOREM 5.2.** Let the conditions of Theorem 3.1 hold. With the window width (5.32) we have

\[
n^{2/(k+4)} \left( \hat{g}_n(x) - g(x) \right) = N(c \chi(x)/f(x) \cdot c^{-1}(\sigma(x)/f(x)) \int K(y) dy),
\]

Note that the asymptotic rate of convergence in distribution is negatively related to the number of regressors. This is typical for nonparametric regression, for the more regressors we have, the more information we ask from the data and thus the more observations we need to get a useful answer.

The result (5.33) has only practical significance if \( n \) is possible to estimate the mean and the variance of the limiting normal distribution involved. As far as the variance is concerned, consistent estimation will appear to be feasible. Regarding the mean, however, the estimation problem is too hard. It would therefore be preferable to get rid of the mean of the limiting normal distribution. We already mentioned a way to do that, namely to choose the window width such that the limit \( \mu \) in (5.27) is zero, but then we also sacrifice some of the speed of convergence. There is, however, another way to get rid of the
asymptotic bias while maintaining the maximal rate of convergence in distribution of order $n^{-2/(k+4)}$, namely by combining the results (5.28) and (5.30). The idea is to use (5.30) for estimating the mean of the limiting normal distribution in (5.28) by subtracting the random function at the left hand side of (5.30) times $\mu$ from the left hand side of (5.28).

**Theorem 5.3.** Let the conditions of Theorem 5.1 hold. Let $\tilde{g}(s)$ be the kernel regression estimator with window width $n^{-1/(k+4)}$ (5.24)

and let $\hat{g}(s)$ be the kernel regression estimator with window width $n^{-\delta/(k+4)}$, with $\delta \in (0,1)$ (5.25).

Denote

$$\delta(s) = \frac{\tilde{g}(s) - n^{-2(\delta - \delta)/(k+4)}\tilde{g}(s)}{(1-n^{-2(1-\delta)/(k+4)})}.$$ (5.26)

Then

$$n^{2/(k+4)}(\tilde{g}(s) - g(s)) \converges to N(0,c^*\beta(s)/f(s))\int K(\cdot)dx \text{ in distr.}$$ (5.27)

This result is related to the generalised jackknife method of Schucany and Sommers (1977) for bias reduction of kernel density estimators.

The rate of convergence is distribution is determined by the rate of convergence of the expectation $q(s)$. If we would choose the kernel $K$ such that $\int xK(x)dx = 0$ and $\int x^2K(x)dx = 0$, then it can be shown that instead of (5.25).

$$\lim_{n \to \infty} n^{-1/2} q(s) = 0.$$ (5.35)
exists and is finite. The asymptotic rate of convergence in distribution then becomes $n^{3/2(k+1)}$ instead of $n^{2/(3k+4)}$. Thus a way to improve the convergence in distribution is to choose a kernel with zero moments up to a particular order $m$. More precisely, following Singh (1981) we define the class $K_m$ of these kernels as follows.

**DEFINITION 5.1.** Let $K_m$ be the class of all bounded Borel measurable real valued functions $K(\cdot)$ on $\mathbb{R}^k$ such that, with $x=(x_1, \ldots, x_k)$, $x_i \in \mathbb{R},$

$$
\int |x|^{1/2 \cdots 1/2 - 1/k} K(x_1, x_2, \ldots, x_k) dx_1 \cdots dx_k = 0 \text{ if } j_1 = j_2 = \cdots = j_{k}=0, \quad 0 \leq i_1 + \cdots + i_k < m,
$$

(5.35)

$$
\int |x|^{1/2} K(x) dx = n \text{ for } i=0 \text{ and } i=m, \quad \int K(x) dx = 1.
$$

With $K \in K_m$ there exists a function $b^*(x)$ such that

(5.40) \[ \lim_{n \to \infty} \gamma^n_b q_i(x) = b^*(x), \]

provided $f(x)$ and $g(x)f(x)$ belong to the class $D_m$.

**DEFINITION 5.2.** Let $D_m$ be the class of all continuous real functions $p$ on $\mathbb{R}^k$ such that the derivatives \( \frac{\partial}{\partial x_1} \cdots \frac{\partial}{\partial x_k} p(x_1, \ldots, x_k), \quad j_1 + \cdots + j_k < m \), are continuous and uniformly bounded for $0 \leq i_1 + \cdots + i_k \leq m$.

Similarly to Theorem 5.3 we now have:

**THEOREM 5.4.** Let assumption 5.1 and the additional conditions $f(x) \in D_m$, $g(x)f(x) \in D_m$, $K \in K_m$ hold, where $m$ is an integer $> 1$. Let $f(x) > 0$. Let $\tilde{g}(x)$ be the kernel regression estimator with window width.

(5.41) \[ \gamma_0 = n^{-1/(2m+1)} \]
and let $g_0(x)$ be the kernel regression estimator with window with

$$g_0 = c_n^{-1/2(2m + k)}$$ with $c \in (0, 1)$. 

Denote

$$g_n(x) = \left( g_n(x) - g_0(x) \right)\left( 1 - (1 - n/m)^{2m + k} \right)^{-1}.$$ 

Then

$$n^{m/(2m + k)}\left( g_n(x) - g(x) \right) \to N(0, c_n^4 \text{E}(g_0(x)/f(x))^2 K(x)^2 dx).$$

The usual asymptotic normality results in parametric regression analysis hold with a rate of convergence in distribution equal to the square root of the number of observations. We see from (5.44) that in the nonparametric regression case this rate can be approached arbitrarily close by increasing $n$. 

Finally we consider estimation of the asymptotic variance in (5.44). Let

$$\sigma^2_n(x, c) = \frac{c_n^4 \text{E}(g_n(x)/f(x))^2 K((x-x_j)/g_0(x))^2 \text{E}(g_n(x))/g_0(x))^2}{(1/n) \text{E}(K((x-x_j)/g_0(x))^2)}$$

with

$$g_0 = c_n^{-1/(2m + k)}.$$ 

It is not too hard to show:

**Theorem 5.5.** Under the conditions of Theorem 5.4,

$$\lim_{n \to \infty} \sigma^2_n(x, c) = c_n^4 \text{E}(g_0(x)/f(x))^2 K(x)^2 dx,$$
\[
\begin{align*}
\eta(x,c) & = n^{\nicefrac{1}{2} (2m-2k)}(\hat{\rho}_c(x,c) - \rho(x,c)) + o(1/n) \quad \text{in distr.} \\
\text{Moreover, for distinct non-random points } x^{(1)}, x^{(2)}, \ldots, x^{(M)} \text{ we have that } \hat{\rho}(x^{(i)},c), \ldots, \hat{\rho}(x^{(M)},c) \text{ are asymptotically independent.}
\end{align*}
\]

On the basis of this result it is now easy to construct confidence bands for \( g(x) \). In particular, the 95% asymptotic confidence interval for \( g(x) \) is

\[
\hat{g}_c(x,c) = \hat{g}_c(x,c) + 1.96 \sqrt{\hat{\sigma}^2_c(x,c)/(n^{\nicefrac{1}{2} (2m-2k)})}.
\]

5.3 Mixed continuous-discrete regressors

We now consider the case where the first \( k_1 \) components of \( x_j \) are continuous and the remaining \( k_2 \) components are discrete. This is the case we shall encounter in our empirical application, where \( k_1 = 1 \) and \( k_2 = 2 \).

**ASSUMPTION 5.2.** Let \( x_j = (x_j^{(1)}, x_j^{(2)})' \in X \times X_0 \), where \( X_1 \) is a \( k_1 \)-dimensional real space and \( X_2 \) is a subset of a \( k_2 \)-dimensional real space. The set \( X_2 \) is such that

\[
\begin{align*}
& (i) \quad s^{(2)} \in X_2 \text{ implies } g(s^{(2)}) = P(s^{(2)}) = x^{(2)}; \\
& (ii) \quad \forall s^{(2)} \in X_2, g(s^{(2)}) = 1; \\
& (iii) \quad \forall s^{(2)} \in X_2, \forall s^{(1)} \in X_1, s \text{ bounded subset of } X, \text{ finite.}
\end{align*}
\]

Let \( x = (x^{(1)}, x^{(2)})' \in X \times X_0 \) and let \( f(x^{(1)}, x^{(2)}) \) be the density of the conditional distribution of \( x_j^{(1)} \) relative to the event \( x_j^{(2)} = x^{(2)} \). For every fixed \( x^{(2)} \in X_2 \) the following holds:

\[
\begin{align*}
& (iv) \quad \forall x^{(1)}, x^{(2)} \in X, \forall x_j^{(1)} \in X_2 \text{, } f(x^{(1)}, x^{(2)}) \text{ is continuous and uniformly bounded on } X_0; \\
& (v) \quad \forall x^{(1)}, x^{(2)} \in X, \forall x_j^{(1)} \in X_2 \text{, } g(x^{(1)}, x^{(2)}) \text{ is uniformly bounded on } X_0; \\
& (vi) \quad \forall x^{(1)}, x^{(2)} \in X, \forall x_j^{(1)} \in X_2 \text{, } g(x^{(1)}, x^{(2)}) \text{ is continuous and uniformly bounded on } X_0.
\end{align*}
\]
Moreover, we now choose the kernel $K(x^{(1)}, x^{(2)})$ and the window width $\gamma_0$ such that with $(x \omega_0 x')^x \in X_\omega X_Y$ and for $x = x_0$,

$$\gamma_0 > 0, \int \gamma_0 > 0, \int K(x, z) dz = 0 \text{ for every } x > 0;$$

$\gamma_0 \psi_m \in K(z, 0) \in K(x, x')$ with $m \geq 2$;

$$\int K(x, 0) \psi_m dz = 1, \int K(x, x') d\lambda = 1.$$ 

Denoting

$$f(x) = f(x^{(1)}, x^{(2)}) = f(x^{(1)}; x^{(2)}) \psi_m(x^2),$$

we have:

**Theorem 5.6.** Under Assumption 1.2 and condition (5.40) the conclusions of Theorems 5.1-5.3 carry over with $k$ replaced by $k_1$, and $\int K(x, 0) \psi_m dz$ replaced by $\int K(x, 0) \psi_m dz$.

The idea behind this result is that the kernel $K(x, x')$ defined in (5.40) acts asymptotically as the kernel $K(x, x')$ where $K(x, x')$ is the indicator function. Consequently, the kernel regression estimator involved is asymptotically equivalent to the one for the continuous case applied to the subsample of regular points $(y_i, x_i^{(1)})$ for which $x_i^{(1)} = x_i^{(2)}$.

See also Section 5.5.

**3.4 The choice of the kernel**

Beneden (1987) proposed the following simple way to construct kernels in $X_\omega X_Y$ for arbitrary $k \geq 1$ and even $n_0 \geq 2$. For $j = \ldots, m/2$ let the $a_j$ and $\sigma_j$ be such that

$$\sum_{j=1}^{m/2} \sigma_j = 1,$$

$$\sum_{j=1}^{m/2} a_j \sigma_j = 0 \text{ for } i = 1, \ldots, (m/2) - 1.$$
In the empirical application under review we have chosen: \( \sigma = \sqrt{3} \) \((j = 1, 2, \ldots)\) : \( m = 8 \). With this choice the \( \theta \) can be solved from (5.51) and (5.52). Moreover, let for \( x \in \mathbb{R}^k \)

\[
K_0(x(i)) = \frac{\theta^{1/2}}{\theta} \exp[-\frac{1}{2}(x - \mu)^T \Delta^{-1}(x - \mu)] \quad \text{for } i = 1, \ldots, m, \quad \text{Eqn. (5.53)}
\]

where \( \Delta \) is a positive definite matrix. It is not hard to verify that \( K_0(\cdot) \in \mathcal{K}_{\text{adm}} \). A suitable kernel satisfying (5.49) can be constructed similarly to (5.53), i.e., let

\[
K_0^*(x(i)) = \frac{\Delta^{1/2}}{\theta} \exp[-\frac{1}{2}(x - \mu)^T \Delta^{-1}(x - \mu)] \quad \text{for } i = 1, \ldots, m, \quad \text{Eqn. (5.54)}
\]

where \( \Omega_1 \) is the inverse of the upper-left \((k_1 \times k_1)\) submatrix of \( \Omega \) and the \( \sigma_j \)'s and \( \theta_j \)'s are the same as in (5.51) and (5.52).

The question now arises how the matrix \( \Omega \) in (5.53) and (5.54) should be specified. A heuristic approach to solve this problem is to specify \( \sigma \) such that certain properties of the true regression function carry over to the estimate \( \hat{g} \). The property considered in Bierens (1987) is the linear translation invariance principle. Suppose we apply a linear translation to \( x \) and the \( x_j \):

\[
x \quad \text{to } (x_j) = P x_j + t_j, \quad x \quad \text{to } (x_j) = P x_j + t_j, \quad \text{Eqn. (5.55)}
\]

where \( P \) is a non-singular \( k \times k \) matrix and \( t \) is a \( k \)-component vector. Then

\[
g(x) = E(y_j | x_j = x) = E(y_j | x_j = x) = g(x), \quad \text{Eqn. (5.56)}
\]

say. However, if we replace the \( x_j \) and \( x \) in (5.43) by \( x_j^* \) and \( x^* \), respectively, and if we leave the kernel \( K \) unchanged, then the resulting kernel regression estimator \( \hat{g}(x^*) \), \( \text{say} \), will in general be unequal to \( \hat{g}(x) \). The only way to accomplish \( \hat{g}(x^*) = \hat{g}(x) \) in all cases (5.55) is to let the kernel be of the form
(5.57) \[ K(x) = \eta(x', \tilde{V}, x) \]

where \( \eta \) is a real function on \( R \) and \( \tilde{V} \) is the sample variance matrix, i.e.,

(5.58) \[ \tilde{V} = \frac{1}{n} \sum_{j=1}^{n} (x_j - \bar{x})(x_j - \bar{x})' \quad \text{with} \quad \bar{x} = \frac{1}{n} \sum_{j=1}^{n} x_j. \]

In particular, if we use kernels of the form (5.53) and (5.54) then we should specify \( \Omega = \tilde{V} \).

Thus, for \( m = 2, 4, 6, \ldots \), we let

(5.59a) \[ \tilde{K}_m(x) = K_m(x | \tilde{V}) \]

in the continuous case and

\[ K_m(x) = K_m(x | V) \]

(5.59b) \[ \text{in the mixed continuous–discrete case. Moreover:} \]

\[ \text{ASSUMPTION 5.5.} \quad \text{Let} \quad \mathbb{E}|x_j|^{4+\delta} < \infty \quad \text{and let the matrix} \quad V = \text{var}(x_j) \quad \text{be non–singular.} \]

Denoting

(5.60a) \[ K_m(x) = K_m(x | V) \]

in the continuous case and

(5.60b) \[ K_m(x) = K_m(x | V) \]

in the mixed continuous–discrete case, we now have:

THEOREM 5.7. Under Assumptions 5.1–5.3 the kernel regression estimator with kernel
(3.59) has the same asymptotic properties (as previously considered) as the kernel regression estimator with kernel (3.60).


5.5 Sample selection

We recall that the data set on which the nonparametric regression results were based is a further subsample of size 1130 from a subsample of size 2000. The latter subsample is a random sample, but the former is obtained by deleting the households with only one parent or adult and the households with one or two persons in the age group 65 or over, and is therefore not a random sample. In this subsection we show how to account for this sample selection.

Let the original random sample be

\[ (\tilde{y}, \tilde{x}_0, \tilde{x}_1, \ldots, \tilde{x}_k, \tilde{z}_n), \]

where \( \tilde{y} \) is the dependent variable, \( \tilde{x}_j \) is a \( k \)-vector of regressors and \( \tilde{z}_j \) is a dummy variable taking the values 0 or 1. In the empirical application under review we have \( N = 2000 \), \( \tilde{z}_j \) is one of the two expenditure categories, \( \tilde{z} = (\tilde{z}_{12}, \tilde{z}_{13})' \) with

- \( \tilde{z}_{12} = \text{net income}, \)
- \( \tilde{z}_{13} = \text{number of children in the age group 0-15}, \)
- \( \tilde{z}_{14} = \text{number of children in the age group 16 or over} \)

and

- \( \tilde{z}_0 = 0 \) for households with only one adult (parent) or with one or two persons in the age group 65 or over,
- \( \tilde{z}_1 = 1 \) for other households.

We now assume:

...
ASSUMPTION 5.4. For this random sample the previous assumptions hold (reading $y_i = \tilde{y}_i$, $x_i = (\tilde{x}_i, \tilde{x}_j)$, $k = k + 1$, $a = N$).

We are interested in estimating the conditional expectation

\begin{equation}
\hat{g}(x) = E[\tilde{y}_i | \tilde{x}_i = x, \tilde{z}_i = 1].
\end{equation}

\text{(5.62)}

Let \( \{\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_n\} \) be a further subsample of size \( n \) corresponding to the data points \( \{\tilde{y}_i, \tilde{x}_i\} \) for which \( \tilde{z}_i = 1 \). Calculate the kernel regression estimator \( \hat{g}_k(x|c) \) and the variance estimator \( \hat{\sigma}^2_k(x|c) \) as if the conditions of Theorem 5.6 hold. Moreover, let \( h_{\{1\}}(s^{(1)}) \) be the conditional density of \( \tilde{z}_i^{(1)} \) relative to the event \( (\tilde{z}_i^{(1)}, \tilde{x}_i^{(1)}) = (x^{(1)}, 1) \in X \times \{1\} \), and let

\begin{equation}
\hat{p}(x^{(2)}) = P(\tilde{z}_i^{(1)} = x^{(1)}, \tilde{z}_i^{(2)} = 1 | \tilde{x}_i = h(x^{(1)}) | x^{(1)}) p(x^{(2)}).
\end{equation}

\text{(5.63)}

THEOREM 5.8. Under Assumption 5.4 the following holds:

(I) For every \( x \) with \( h(x) > 0 \) and each constant \( c > 0 \),

\begin{equation}
N^{\max\{2m-k+1\}} |\hat{g}_k(x|c) - g(x) - N[0, \hat{\sigma}^2_k(x|c)] | \text{ in distr.}
\end{equation}

\text{(5.64)}

\text{where}

\begin{equation}
\hat{\sigma}^2_k(x|c) = \lim_{\rightarrow 0} \hat{\sigma}^2_k(\hat{y}_i|c).
\end{equation}

\text{(5.65)}

(II) Let \( x^{(1)}, \ldots, x^{(M)} \) be distinct points for which \( h(x^{(1)}) > 0 \) and let

\begin{equation}
\hat{\delta}_k(x|c) = N^{\max\{2m-k+1\}} |\hat{g}_k(x|c) - g(x)| \hat{\alpha}_k(\hat{y}_i|c)
\end{equation}

\text{(5.66)}

Then \( \hat{\delta}_k(x^{(1)}|c), \ldots, \hat{\delta}_k(x^{(M)}|c) \) are asymptotically independent \( N(0,1) \) variables.
Proof. Let \( K_a(x) \) be the kernel (5.60b) calculated on the basis of the subsample of size \( n \).

Define

\[
\hat{a}_N(x, s) = K_a(s) \cdot I(n = 0),
\]

where \( I(\cdot) \) is the indicator function. Defining \( \hat{g}_N(x, s|c) \) and \( \hat{a}_N(x, s|c) \) similarly to (5.43)

and (5.45), respectively, with kernel (5.68), the results in Theorem 5.6 for the random sample of size \( N \) go through. The theorem now follows from the fact that

\[
\hat{g}_N(x|c) = \hat{g}_N(x, s|c), \quad \hat{a}_N(x|c) = \hat{a}_N(x, s|c).
\]

5.6 Choosing the constant \( c \)

In Bieres (1987) it is advocated to choose the constant \( c \) of the window width by cross-validation. See Marron (1981) for a review of the literature on cross-validation and other methods of adaptive window width selection. Also, see Gasser (1973), Stone (1974) and Wahba and Wold (1976) and Li (1987). In the cross-validation approach each \( y_j \) in the sample of size \( n \) is predicted by the kernel regression estimator based on the remaining \( n-1 \) observations. Thus let \( \hat{g}_N(x|c) \) be the kernel regression estimator based on the sample with the \( j \)-th observation left out. Then \( c \) is determined by minimizing the objective function

\[
Q(c) = \sum_{j=1}^{n} \left( y_j - \hat{g}_N(x|c) \right)^2
\]

over an interval \([c_1, c_2]\), \( 0 < c_1 < c_2 < n \). A drawback of this approach is that the resulting estimated constant \( c \), say, depends on the same sample as the kernel regression function estimator. Consequently, \( \hat{g}(x|c) \) with \( c \) fixed is not independent of \( c \) and hence the asymptotic normality results for \( g(x|c) \) may not hold for \( \hat{g}(x|c) \). Therefore we have used the smaller random subsample of size 850 for estimating \( c \) by cross-validation. Then \( c \) is independent of the kernel regression estimator \( \hat{g}(x|c) \) based on the random subsample of size 2000 and therefore all the asymptotic normality results carry over to \( \hat{g}(x|c) \).
The resulting cross-validated \( c \), however, appeared to be too large, by which the kernel regression estimator became almost constant. Therefore we have conducted various experiments with alternative values of \( c \), still confining the analysis to the smaller subsample. It appeared that the best choice for \( c \) was \( c = 2 \); best in the sense that for this \( c \) the kernel regression estimate was sufficiently smooth without being flat. Using \( c = 2 \) the nonparametric regression analysis has been further conducted on the basis of the larger subsample of size 2000.

6. SPECIFYING HOUSEHOLD EXPENDITURE FUNCTIONS BY NONPARAMETRIC REGRESSION

6.1 Nonparametric regression results

The nonparametric regression results for the Engel functions under review are displayed in Figures 1 to 16. The first 8 figures show the kernel regression estimator (the solid line) for expenditures on food, clothing and footwear, where \( x_1 \) (net income) runs from 16,000 to 65,000 guilders. In some cases the income range is smaller, due to lack of observations in the low and high income range. The scale of the figures is linear on both axes. Each figure corresponds to a household type \((x_2,x_3)\), where \( x_2 \) is the number of children in the age group 0–15 and \( x_3 \) is the number of children in the age group 16 or over. We only show the nonparametric results for households with \( 0 \leq x_2 \leq 3 \) and \( 0 \leq x_3 \leq 1 \), or other households are too rare. The dotted lines are the 95% confidence bands. The other 8 figures show the nonparametric results for other expenditures.

We recall that the kernel regression estimate is the one in Theorem 3.8 with kernel of the type \((5.59b)\) with \( \sigma = \sqrt{2} \) and \( m = 8 \) (cf. \((5.51)\) and \((5.53)\)) and window width with constant \( c = 2 \). Thus the asymptotic rate of convergence of the kernel estimator in \((5.64)\) is \( N^{-1/2} \). The 95% confidence bands are based on \((5.67)\). Observe that the 95% confidence band becomes wider in the low and high income range, due to lack of observations. Also the deviation from linearity of the kernel estimator in various figures (especially in Figures 8 and 16 and in the low and high income ranges in other figures) is probably due to lack of observations.
Figure 3

Figure 4
Figure 13: Other expenditures: Household type (0,1) and 95% confidence band.

Figure 14: Other expenditures: Household type (1,1) and 95% confidence band.
Figure 15

Figure 16
In the next section we shall interpret these results. In particular we shall pay attention to the question how to translate them to parametric specifications.

6.2 Parametrisation of the nonparametric regression results

The nonparametric regression results for expenditures on food, clothing and footwear indicate that the income–expenditure relationships involved are almost linear: for each household type it is possible to draw a straight line almost entirely inside the 95% confidence band. The estimated Engel curves are only curved in the low and high income ranges. These curved parts, however, need not be significantly different from a straight line, as indicated by the 95% confidence bands, as, in nonparametric regression analyses, estimation errors manifest themselves in the form of bumps on the estimated regression curve. Thus the nonparametric regression results indicate that over the income range 16,000–65,000 the Engel curves involved are linear. The same applies to other expenditures. However, although straight lines can be fitted within all confidence bands, it is possible that after flexible parametrisation of the Engel curves the bending parts appear significant, because a correct parametrisation gives a regression function a "backbone" which will narrow the confidence bands.

The curvature of the kernel estimates is either a stretched reversed S–shape or a U–shape. This suggests to specify the Engel curves involved by third–order polynomials in income. In order to check this specification we have approximated the kernel regression function estimators for each household type by third–order polynomials in net income, by regressing the kernel estimator on \( x_1, x_1^2 \) and \( x_1^3 \) for grid points in the interval 16,000–65,000. It appears that for each household type this polynomial approximation fits very well in the 95% confidence band, which indicates that a third–order polynomial is a suitable functional form for the household Engel function under review. The third–order polynomial approximations are in fact so close that they hardly can be distinguished from the corresponding kernel estimators. Therefore we cannot show them in the figures.

The parameters of the third–order polynomials can be made dependent of the number of children in the household by using the following dummy variables:
The parametrisation of the nonparametric regression results then becomes:

\[ y_{ij} = \alpha_0 + \beta_0 x_{ij} + \delta x_{ij}^2 + \epsilon_{ij}, \]

\[ + \epsilon_{ij}(\mu_i + \lambda_i + \alpha_i x_{ij} + \beta_i x_{ij}^2) + \epsilon_{ij}, \quad i=1,2. \]

We have used a subsample of the original sample of size 2859 to estimate the parameters involved by OLS. This subsample consists of all households of the type \((x_0, x_1)\) with \(x_1 \leq 3\) and \(x_2 \leq 1\), net income \(x_3\) in the range 16,000–65,000, and two parents both younger than 65. The size of this subsample is 1302. Since the errors are likely heteroscedastic, we have calculated the variance matrix of the OLS estimates of the parameters of model (6.2) according to the approach of White (1980). Moreover, in order to enhance the numerical stability of the OLS estimates and the tests we have scaled down the income variable \(x_{ij}\) into the interval \((0,1)\) by dividing it by 65,000, and all calculations have been conducted in double precision. The estimation results involved are presented in Table 1.

The test of the linearity hypothesis amounts to testing the null hypothesis

\[ H_0: \beta_m = 0 \quad \text{for } m=0,1,2,3. \]

The test statistics of the Wald test involved are 12.56 for expenditures on food, clothing and footwear and 10.36 for other expenditures. Under the null hypothesis these test statistics are asymptotically \(\chi^2_1\) distributed, hence the linearity hypothesis cannot be rejected at any reasonable significance level.

Next we have tested whether the linearity hypothesis holds with constant slope. Thus the null hypothesis to be tested is now:

\[ H_0: \beta_m = 0 \quad \text{for } m=1,2,3,4, \quad \text{and } \mu_i = 0 \quad \text{for } m=0,1,2,3,4. \]
<table>
<thead>
<tr>
<th></th>
<th>Food, Clothing &amp; Footwear</th>
<th>Other expenditures</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>estimate</td>
<td>t-value</td>
<td>estimate</td>
<td>t-value</td>
</tr>
<tr>
<td>a0</td>
<td>5852.0</td>
<td>1.0070</td>
<td>9159.0</td>
<td>0.2762</td>
</tr>
<tr>
<td>a1</td>
<td>8691.0</td>
<td>1.0730</td>
<td>6739.0</td>
<td>0.3039</td>
</tr>
<tr>
<td>a2</td>
<td>-436.3</td>
<td>-0.8556</td>
<td>-436.3</td>
<td>-0.8556</td>
</tr>
<tr>
<td>a3</td>
<td>6984.0</td>
<td>0.6717</td>
<td>5900.0</td>
<td>1.8680</td>
</tr>
<tr>
<td>a4</td>
<td>-13310.0</td>
<td>-1.4164</td>
<td>-26670.0</td>
<td>-0.9795</td>
</tr>
<tr>
<td>b0</td>
<td>2830.0</td>
<td>0.1044</td>
<td>18480.0</td>
<td>0.1762</td>
</tr>
<tr>
<td>b1</td>
<td>-48020.0</td>
<td>-1.1429</td>
<td>-28140.0</td>
<td>-0.2724</td>
</tr>
<tr>
<td>b2</td>
<td>12680.0</td>
<td>0.3152</td>
<td>35160.0</td>
<td>0.3542</td>
</tr>
<tr>
<td>b3</td>
<td>-19920.0</td>
<td>-0.3679</td>
<td>-29120.0</td>
<td>-1.7490</td>
</tr>
<tr>
<td>b4</td>
<td>83220.0</td>
<td>1.7799</td>
<td>10020.0</td>
<td>1.1570</td>
</tr>
<tr>
<td>c0</td>
<td>89652.0</td>
<td>0.1918</td>
<td>33735.0</td>
<td>0.2531</td>
</tr>
<tr>
<td>c1</td>
<td>19730.0</td>
<td>1.3190</td>
<td>54220.0</td>
<td>0.2086</td>
</tr>
<tr>
<td>c2</td>
<td>-24630.0</td>
<td>-0.2676</td>
<td>-36160.0</td>
<td>-0.2022</td>
</tr>
<tr>
<td>c3</td>
<td>28330.0</td>
<td>0.3160</td>
<td>44710.0</td>
<td>1.6790</td>
</tr>
<tr>
<td>c4</td>
<td>-13830.0</td>
<td>-1.8020</td>
<td>-29040.0</td>
<td>-1.2700</td>
</tr>
<tr>
<td>c5</td>
<td>-6220.0</td>
<td>-0.2470</td>
<td>-12390.0</td>
<td>0.2728</td>
</tr>
<tr>
<td>c6</td>
<td>-53850.0</td>
<td>-1.4740</td>
<td>-27410.0</td>
<td>-0.3482</td>
</tr>
<tr>
<td>c7</td>
<td>15290.0</td>
<td>0.4240</td>
<td>13480.0</td>
<td>0.3888</td>
</tr>
<tr>
<td>c8</td>
<td>-13660.0</td>
<td>-0.3930</td>
<td>-21590.0</td>
<td>-1.5390</td>
</tr>
<tr>
<td>c9</td>
<td>71660.0</td>
<td>1.7880</td>
<td>109860.0</td>
<td>1.3650</td>
</tr>
</tbody>
</table>

SE 3328.0 7530.0
R² 0.1641 0.3198
n 1502 1592

The test statistics of the Wald test are 16.13 for expenditures on food, clothing and
footwear and 15.93 for other expenditures. Under H0 these test statistics are
asymptotically χ²-distributed, and consequently also this null hypothesis cannot be
rejected. Thus the model reduces to:

\[ y_{ij} = a_0 + a_1 x_{ij} + a_2 x_{ij} + a_3 x_{ij} + a_4 x_{ij} + b_{ij} + x_{ij} = 1,2. \]

Furthermore, we have tested whether this model can be written as a linear regression
model with explanatory variables \( x_{ij} \) and \( x_{ij} \). This simplification corresponds to the
following hypothesis:
The Wald statistics involved are 16.51 for food, clothing and footwear and 20.09 for other expenditures. Under the null these statistics are asymptotically χ² distributed and therefore we cannot reject the null hypothesis.

Finally we have tested whether the Engel curves are linear and independent of the household size. This hypothesis corresponds to:

\[ \text{H}_5: \alpha_m = \beta_m = 0 \text{ for } m = 1, 2, 3, 4; \delta_m = \epsilon_m = 0 \text{ for } m = 0, 1, 2, 3, 4. \]

The Wald statistics are 156.2 for food, clothing and footwear and 20.14 for other expenditures. Since \( P[\chi^2_{15} > 156.2] \) cannot be distinguished from zero we have to reject this null hypothesis for expenditures on food, clothing and footwear, while the hypothesis involved cannot be rejected for other expenditures.

Table 2 summarizes the test results. Note that the tests involved are not independent. From a formal point of view we should therefore not re-estimate the model.
after each test as otherwise the type I errors may accumulate. Nonetheless we have
checked the final conclusions by conducting a similar sequence of tests starting from the
linear model with slope and intercept depending on family size, and the linear model
with constant slope and intercept depending on family size, respectively. These tests lead
to the same conclusion as before, namely that the household Engel function for
expenditures on food, clothing and footwear is a linear function in net income (x1), the
number of children is the age group 0-11 (x3) and the number of children in the age group
16 or over (x5), while the household Engel function for other expenditures is a linear
function in net income only.

6.3 Final models

The simplification of the polynomial model suggested by the test results in Section
6.2 now lead to the following models:

\[
(y_1) \quad \text{Food, clothing and footwear:} \\
\beta_1 = 0.00736 x_1 + 780.5 x_2 + 21.36 x_3 + 562 \quad (11.41) \quad (8.23) \quad (8.28) \quad (17.5) \\
E^2 = 0.1580; SE = 3129; CCMT = 2.316; n = 1502
\]

\[
(y_2) \quad \text{Other expenditures:} \\
\beta_2 = 0.5115 x_1 + 6114 \quad (23.4) \quad (7.79) \\
R^2 = 0.3104; SE = 7637; CCMT = 0.0081; n = 1507
\]

with the \( t \)-values in parentheses.

The statistics CCMT are the chi-square (1) statistics of Breuer’s (1889) consistent
conditional moment test of functional form. See the Appendix. These test statistics
indicate that the null hypothesis that models (6.8) and (6.9) represent the conditional
expectation of \( y_1 \) and \( y_2 \), respectively, relative to \( x_1 \), \( x_2 \) and \( x_3 \), cannot be rejected.
Consequently, it is of no use to investigate other functional forms, in particular the popular
Working-Leser form. Since the test involved is consistent and the sample size is quite
large we certainly would have detected such a severe deviation from linearity.
The correlation coefficient of the errors \( u_1 \) and \( u_2 \) of models (6.8) and (6.9), respectively, is 0.249. Moreover, denoting
\[
b = (b_0, b_1, b_2)'
\]
and \( \delta \) the vector of true parameters, the asymptotic variance matrix of \( \hat{b} \) is:
\[
(\hat{b} - b)'
\]
and the corresponding correlation matrix is:
\[
r_{0.000} = -0.1454
-0.2679
-0.8801
0.2281
0.2206
0.0467
0.0503
0.2211
-0.9885
1.0000
1.0000
Note that model (6.8) can be written as
\[
y_j(x_{2j}, x_3) = 3813.5 + 0.09736 x_j + \eta(x_{2j}, x_3)
\]
where
\[
\eta(x_{2j}, x_3) = 2 + 0.2774 x_1 + 0.7599 x_j,
\]
The t-values involved follow from (6.8), (6.9) and (6.10).
It is tempting to interpret (6.12) as the adult equivalence scale. Thus, as far as food, nothing and footwear is concerned a child under 16 counts for about 28% of an adult.
and a dependent child of 16 or over counts for about 76% of an adult, in a household with two parents and net income in the range 16,000–65,000. However, since price effects could not be taken into account it is impossible to recover the utility function and cost function from (6.8) and (6.9). Therefore the "adult equivalence scale" (6.12) cannot be used for household welfare comparison.

It should be stressed that the lack of impact of household size on the overall expenditures does not imply that there is no impact at all. It is likely that the extra expenditures due to children will be covered by substitution within the same expenditure category. For example the extra expenditures on housing may be covered by a cheaper car rather than a new one, etc. etc.

As said before, the household Engel functions considered in the literature usually relate expenditures on various commodities to total expenditure rather than to income, in order to impose the budget constraint and to interpret the models in terms of utility theory. However, in view of the argument in Section 3 we should instead relate the two expenditure categories to the conditional expectation of total expenditure rather than total expenditure itself. We can put the above models in this form by solving equation (6.9) to \( x_1 \) and substituting the result for \( x_1 \) in equation (6.9). This yields, after some elementary calculations:

\[
y = 0.1599 \hat{y} + 55.7 x_2 + 1796 x_3 + 3749 \tag{6.13}
\]

where \( y = y_1 + y_2 \). This model can also be written as:

\[
y = \frac{n(x_0, x_3)}{1874.5 + 0.1599 \hat{y} / n(x_0, x_3)} \tag{6.14}
\]

with

\[
h(x_0, x_3) = 2 + 0.3498 x_2 + 0.0482 x_3 \tag{6.15}
\]

The \( t \)-values involved have been calculated on the basis of (6.10).
Again it is tempting (even more than in the case (6.12)) to interpret (6.15) as an adult equivalence scale: a child under 16 counts for about 35% of an adult, and a child of 16 or over counts for about 96% of an adult. The latter percentage is not significantly different from 100%; the t-test statistic of the null hypothesis involved is -0.3019. These results look more plausible than the corresponding results of (6.12). However, also now we cannot use (6.15) for welfare comparison, as the utility and cost functions cannot be recovered from (6.15). Only if we would follow the original Engel approach and assume that households have the same welfare level if their budget shares for food, clothing and footwear are the same can we interpret (6.15) as an adult equivalence scale.

7. CONCLUDING REMARKS

One may object against our approach and our results that there is no relationship at all with economic theory and that the above results cannot be interpreted in terms of utility theory. The latter is due to the fact that by working with a single survey no price effects can be taken into account. Moreover, even if we would be able to determine the corresponding cost function welfare comparison with this model is disputable, as the fit of the models (6.8) and (6.9) is quite low. Cf. Section 3. As to the former complaint, we have indeed worked the other way around, i.e., we started with analyzing the data in order to determine the model rather than first setting up the model in order to analyze the data. One should bear in mind, however, that the traditional econometric approach assumes very restrictive household behavior, in particular the implicit assumption that all households face exactly the same utility function. This will unlikely be the case in reality. Our approach focuses on actual household behavior only.

As to our empirical finding, it is apparent that without imposing as a priori functional form, our data set does not contain enough information to distinguish the shape of our household Engel curve from simple linear functions. This may cast doubts about the nonlinear Engel curves found in the literature; the nonlinearity might have been put in. An exception is Blundell and Ray (1984). They tested the linear Engel curve specification into a class of nonlinear Engel curves, using a pooled cross-section consisting of 140 observations on annual time series for the period 1968-1979, and quarterly time series
covering the period 1955–I to 1980–II. Thus Blundell and Ray could also take price effects into account. They rejected the linear specification.

ACKNOWLEDGEMENT

The helpful comments of Wim Groot, Joop Hartog, Arie Kapteyn, Timo Terasingi and three referees are gratefully acknowledged. This paper was presented at the Econometric Society European Meeting 1989 in Munich.

REFERENCES


**APPENDIX: The Consistent Conditional Moment Test**

Bierens (1989) proposed a consistent conditional moment test of functional form of the nonlinear regression model.
This test is consistent against all alternatives, i.e., any deviation of the true model will be detected when the sample size \( n \) goes to infinity. The model is true if the functional form \( f(x, \theta) \) represents the conditional expectation of \( y_j \) relative to the vector \( x_j = (x_{ij}, \ldots, x_{in})' \) of regressors:

\[
(A.2) \quad H_0: P[E(y_j | x_j) = f(x_j, \theta_0)] = 1.
\]

where

\[
(A.3) \quad \delta_j = \arg \min \{E(y_j - f(x_j, \theta))^2 \}.
\]

This is the null hypothesis to be tested. The alternative hypothesis is that the null is false:

\[
(A.4) \quad H_0: P[E(y_j | x_j) = f(x_j, \theta_0)] < 1.
\]

In the empirical application under review we have

\[
(A.5) \quad f(x_j, \theta) = \theta_0 x_{ij} + \theta_1 x_{i2} + \theta_2 x_{i3} + \theta_3 x_{i4} + \theta_4,
\]

with \( x_{ij} \) net income, \( x_{ij} \) the number of children younger than 16 and \( x_{i3} \) the number of children of 16 or over. For the case \( y_j \) other expenditures, we base the test not on (6.9) but on the augmented model:

\[
(A.7) \quad y_j = 0.5117 x_{i1} + 19.10 x_{i3} + 19.44 x_{i4} + 5860, \quad (23.0) \quad (0.962) \quad (0.028) \quad (7.20).
\]

\( R^2 = 0.3108; SE = 7940, n = 1502. \)

Cf. (e.g.)
The test statistic of Bierens' consistent conditional moment test can be constructed as follows. Let \( T \) be a hypercube in \( \mathbb{R}^3 \) not containing the origin and let for \( t \in T \)

\[
(w_{ij}(t)) = \prod_{h=1}^{n} \exp[p((x_{ij} - \bar{x}_j)/s_j)]
\]

where \( x_i \) and \( s_i \) are the sample means and the sample standard error of the \( x_{ij} \), respectively, and \( p \) is a continuously differentiable bounded one-to-one mapping from \( \mathbb{R} \) into \( \mathbb{R} \) with uniformly bounded first derivative. In the empirical application under review we have chosen

\[
T = [x_{ij}(1, 1); s(x) = \tan^{-1}(x/2)].
\]

Next, let \( u_j \) be the least squares residuals and denote

\[
\tilde{W}(t) = n[M(t)^{0.5}/s(t)],
\]

where

\[
\tilde{M}(t) = (1/n)\sum_{i=1}^{n} s_i w(x_{ij}),
\]

\[
\tilde{s}(t) = (1/n)\sum_{i=1}^{n} [w(x_{ij}) - \mu(t)] - \tilde{M}(t)^{0.5} (\partial/\partial \tilde{w})(\tilde{x}_j, \tilde{y})^{0.5},
\]

\[
\tilde{u}(t) = (1/n)\sum_{i=1}^{n} \left( (\partial/\partial \tilde{w})(\tilde{x}_j, \tilde{y})^{0.5} \right) w(x_{ij}),
\]

\[
\tilde{A} = (1/n)\sum_{i=1}^{n} (\partial/\partial \tilde{w})(\tilde{x}_j, \tilde{y})^{0.5} (\tilde{w}(x_{ij})^{0.5}).
\]

Now choose independently of the data generating process real numbers \( \gamma > 0, \rho \in (0,1), \) and a point \( t_0 \in T \). Moreover, choose a sequence of positive integers \( n_k \) converging to infinity with \( n_k \), and choose a sequence \( \{t_i\} \) such that \( \{t_1, t_2, t_3, \ldots\} \) is dense in \( T \). We may choose the \( t_i \)'s \( (i=0,1,2,\ldots) \) randomly from a continuous distribution with density having
support $T$. Let

$$i = \arg\max_{t \in \{1, \ldots, T\}} \tilde{W}(t)$$

and let

$$\mathcal{T} = t_i \text{ if } \tilde{W}(t_i) - \tilde{W}(t_0) > \frac{m^0}{2m^0}$$

Furthermore, denote

$$\eta(t) = \lim_{n \to \infty} \frac{\tilde{W}(t)}{n}$$

Then under $H_0$,

$$\tilde{W}(t)/t \xrightarrow{d} \chi^2$$

whereas under $H_1$, $\eta(t) > 0$ for all $i$ except in a non-dense set with Lebesgue measure zero,

and

$$\tilde{W}(t)/t \xrightarrow{a.s.} \sup_{\eta > 0} \eta(t)$$

Cf. Bierens (1989, Theorems 4 and 5). Thus $\tilde{W}(t)$ is the test statistic of Bierens' consistent conditional moment test, with null distribution $\chi^2$.

In the empirical application under review we have chosen the $t_i$ randomly from the uniform distribution on $T$, and

$$K_0 = \lceil 8 + 0.05n \rceil; \gamma = 0.5; \rho = 0.5$$

This yield for $n=1500$, $K_0 = 75$, $m^0 = 19.38$. 
<table>
<thead>
<tr>
<th>$t_i$</th>
<th>$t_j$</th>
<th>$t_k$</th>
<th>$W(t_i)$</th>
<th>$t_1$</th>
<th>$t_2$</th>
<th>$t_3$</th>
<th>$W(t_i)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3.303</td>
<td>3.476</td>
<td>4.375</td>
<td>2.316</td>
<td>1.382</td>
<td>1.988</td>
<td>4.289</td>
</tr>
<tr>
<td>0</td>
<td>3.303</td>
<td>3.476</td>
<td>4.375</td>
<td>2.316</td>
<td>4.666</td>
<td>1.792</td>
<td>2.796</td>
</tr>
<tr>
<td>1</td>
<td>3.191</td>
<td>3.023</td>
<td>2.071</td>
<td>1.897</td>
<td>5.192</td>
<td>4.119</td>
<td>1.535</td>
</tr>
<tr>
<td>2</td>
<td>1.456</td>
<td>2.831</td>
<td>3.125</td>
<td>1.619</td>
<td>3.984</td>
<td>4.508</td>
<td>4.421</td>
</tr>
<tr>
<td>3</td>
<td>7.866</td>
<td>1.904</td>
<td>4.768</td>
<td>1.240</td>
<td>11.256</td>
<td>1.163</td>
<td>2.281</td>
</tr>
<tr>
<td>4</td>
<td>2.567</td>
<td>1.766</td>
<td>3.908</td>
<td>1.278</td>
<td>1.125</td>
<td>4.109</td>
<td>2.955</td>
</tr>
<tr>
<td>5</td>
<td>1.435</td>
<td>4.922</td>
<td>3.717</td>
<td>1.612</td>
<td>4.077</td>
<td>4.859</td>
<td>1.850</td>
</tr>
<tr>
<td>6</td>
<td>1.010</td>
<td>5.367</td>
<td>6.360</td>
<td>0.992</td>
<td>7.580</td>
<td>3.583</td>
<td>4.807</td>
</tr>
<tr>
<td>7</td>
<td>8.298</td>
<td>4.390</td>
<td>1.823</td>
<td>2.391</td>
<td>1.927</td>
<td>1.420</td>
<td>2.256</td>
</tr>
<tr>
<td>8</td>
<td>1.392</td>
<td>3.816</td>
<td>2.009</td>
<td>2.304</td>
<td>2.051</td>
<td>3.690</td>
<td>3.481</td>
</tr>
<tr>
<td>9</td>
<td>3.420</td>
<td>3.355</td>
<td>2.657</td>
<td>1.975</td>
<td>2.317</td>
<td>3.447</td>
<td>2.985</td>
</tr>
<tr>
<td>9</td>
<td>2.138</td>
<td>1.488</td>
<td>3.826</td>
<td>1.648</td>
<td>2.365</td>
<td>2.888</td>
<td>2.558</td>
</tr>
<tr>
<td>10</td>
<td>2.823</td>
<td>2.653</td>
<td>3.052</td>
<td>1.539</td>
<td>2.476</td>
<td>3.223</td>
<td>1.846</td>
</tr>
<tr>
<td>11</td>
<td>2.055</td>
<td>2.451</td>
<td>3.921</td>
<td>1.537</td>
<td>2.169</td>
<td>3.007</td>
<td>5.810</td>
</tr>
<tr>
<td>12</td>
<td>3.330</td>
<td>1.788</td>
<td>2.105</td>
<td>2.332</td>
<td>2.105</td>
<td>2.909</td>
<td>2.586</td>
</tr>
<tr>
<td>13</td>
<td>2.779</td>
<td>4.448</td>
<td>2.848</td>
<td>2.448</td>
<td>3.418</td>
<td>4.439</td>
<td>1.767</td>
</tr>
<tr>
<td>14</td>
<td>3.171</td>
<td>3.417</td>
<td>4.322</td>
<td>1.188</td>
<td>3.488</td>
<td>2.317</td>
<td>2.014</td>
</tr>
<tr>
<td>15</td>
<td>4.093</td>
<td>1.892</td>
<td>2.766</td>
<td>1.712</td>
<td>3.517</td>
<td>3.672</td>
<td>4.972</td>
</tr>
<tr>
<td>16</td>
<td>3.960</td>
<td>2.252</td>
<td>1.765</td>
<td>1.929</td>
<td>3.056</td>
<td>1.142</td>
<td>4.767</td>
</tr>
<tr>
<td>17</td>
<td>3.750</td>
<td>1.322</td>
<td>2.380</td>
<td>2.359</td>
<td>2.567</td>
<td>4.631</td>
<td>1.277</td>
</tr>
<tr>
<td>18</td>
<td>2.519</td>
<td>1.627</td>
<td>1.824</td>
<td>1.777</td>
<td>2.927</td>
<td>4.437</td>
<td>2.952</td>
</tr>
<tr>
<td>19</td>
<td>4.098</td>
<td>2.236</td>
<td>2.205</td>
<td>2.077</td>
<td>3.577</td>
<td>4.646</td>
<td>2.710</td>
</tr>
<tr>
<td>20</td>
<td>1.534</td>
<td>1.411</td>
<td>1.433</td>
<td>1.728</td>
<td>4.308</td>
<td>2.211</td>
<td>6.659</td>
</tr>
<tr>
<td>21</td>
<td>3.769</td>
<td>4.627</td>
<td>1.397</td>
<td>2.266</td>
<td>4.289</td>
<td>2.554</td>
<td>2.068</td>
</tr>
<tr>
<td>22</td>
<td>2.043</td>
<td>2.379</td>
<td>4.865</td>
<td>1.301</td>
<td>4.091</td>
<td>2.459</td>
<td>2.537</td>
</tr>
<tr>
<td>23</td>
<td>2.826</td>
<td>1.850</td>
<td>1.932</td>
<td>2.280</td>
<td>4.509</td>
<td>2.452</td>
<td>1.590</td>
</tr>
<tr>
<td>24</td>
<td>3.420</td>
<td>4.020</td>
<td>3.017</td>
<td>2.792</td>
<td>5.396</td>
<td>3.096</td>
<td>1.808</td>
</tr>
<tr>
<td>25</td>
<td>3.420</td>
<td>4.669</td>
<td>2.986</td>
<td>1.978</td>
<td>5.185</td>
<td>3.605</td>
<td>3.203</td>
</tr>
<tr>
<td>26</td>
<td>3.560</td>
<td>3.228</td>
<td>2.371</td>
<td>1.648</td>
<td>5.672</td>
<td>1.539</td>
<td>3.292</td>
</tr>
<tr>
<td>27</td>
<td>2.717</td>
<td>2.547</td>
<td>3.055</td>
<td>1.720</td>
<td>4.253</td>
<td>4.681</td>
<td>4.584</td>
</tr>
<tr>
<td>29</td>
<td>4.301</td>
<td>1.686</td>
<td>1.793</td>
<td>1.796</td>
<td>6.551</td>
<td>1.809</td>
<td>3.571</td>
</tr>
<tr>
<td>30</td>
<td>4.701</td>
<td>1.460</td>
<td>1.392</td>
<td>2.499</td>
<td>6.732</td>
<td>2.284</td>
<td>2.870</td>
</tr>
<tr>
<td>31</td>
<td>4.340</td>
<td>1.532</td>
<td>1.875</td>
<td>1.854</td>
<td>7.890</td>
<td>3.934</td>
<td>4.768</td>
</tr>
<tr>
<td>32</td>
<td>4.886</td>
<td>2.663</td>
<td>1.339</td>
<td>1.691</td>
<td>7.419</td>
<td>3.729</td>
<td>2.952</td>
</tr>
<tr>
<td>33</td>
<td>7.918</td>
<td>3.814</td>
<td>4.292</td>
<td>1.383</td>
<td>7.396</td>
<td>3.541</td>
<td>2.171</td>
</tr>
</tbody>
</table>

The values of the random functions $W(t)$ for $t_{1}, t_{2}, \ldots, t_{34}$ and the two expenditure categories are given in tables A1 and A2. From these tables it follows that the costs statistics of the conditional moment tests of models (6.8) and (A.7) are 2.316 and 0.006, respectively, hence the null hypothesis that the models are linear in $x$, $x_{1}$ and $x_{2}$ cannot be rejected at any reasonable significant level.
<table>
<thead>
<tr>
<th></th>
<th>( t_{11} )</th>
<th>( t_{12} )</th>
<th>( t_{13} )</th>
<th>( W(t_{11}) )</th>
<th>( t_{11} )</th>
<th>( t_{12} )</th>
<th>( t_{13} )</th>
<th>( W(t_{11}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4.582</td>
<td>1.313</td>
<td>4.568</td>
<td>0.001</td>
<td>1</td>
<td>3.645</td>
<td>3.002</td>
<td>3.430</td>
</tr>
<tr>
<td>2</td>
<td>2.083</td>
<td>2.305</td>
<td>4.717</td>
<td>4.514</td>
<td>3</td>
<td>2.308</td>
<td>2.984</td>
<td>3.188</td>
</tr>
<tr>
<td>4</td>
<td>4.283</td>
<td>1.830</td>
<td>1.274</td>
<td>0.063</td>
<td>5</td>
<td>1.966</td>
<td>4.559</td>
<td>3.348</td>
</tr>
<tr>
<td>6</td>
<td>2.374</td>
<td>1.889</td>
<td>2.780</td>
<td>2.035</td>
<td>7</td>
<td>1.105</td>
<td>3.112</td>
<td>1.677</td>
</tr>
<tr>
<td>10</td>
<td>1.007</td>
<td>2.965</td>
<td>3.488</td>
<td>4.957</td>
<td>11</td>
<td>1.907</td>
<td>3.054</td>
<td>1.128</td>
</tr>
<tr>
<td>12</td>
<td>2.728</td>
<td>2.188</td>
<td>2.488</td>
<td>2.113</td>
<td>13</td>
<td>2.541</td>
<td>2.878</td>
<td>3.111</td>
</tr>
<tr>
<td>16</td>
<td>3.046</td>
<td>4.769</td>
<td>1.754</td>
<td>2.560</td>
<td>17</td>
<td>4.128</td>
<td>2.023</td>
<td>1.982</td>
</tr>
<tr>
<td>18</td>
<td>1.322</td>
<td>2.334</td>
<td>2.860</td>
<td>4.221</td>
<td>19</td>
<td>3.953</td>
<td>3.471</td>
<td>4.932</td>
</tr>
<tr>
<td>22</td>
<td>1.312</td>
<td>2.566</td>
<td>1.033</td>
<td>3.289</td>
<td>23</td>
<td>2.336</td>
<td>2.982</td>
<td>2.733</td>
</tr>
<tr>
<td>24</td>
<td>3.089</td>
<td>0.948</td>
<td>4.312</td>
<td>3.827</td>
<td>25</td>
<td>1.127</td>
<td>1.374</td>
<td>1.848</td>
</tr>
<tr>
<td>26</td>
<td>1.162</td>
<td>3.914</td>
<td>1.776</td>
<td>5.963</td>
<td>27</td>
<td>4.716</td>
<td>2.778</td>
<td>2.972</td>
</tr>
<tr>
<td>28</td>
<td>2.691</td>
<td>2.630</td>
<td>1.717</td>
<td>2.406</td>
<td>29</td>
<td>2.386</td>
<td>2.429</td>
<td>1.601</td>
</tr>
<tr>
<td>32</td>
<td>1.480</td>
<td>4.554</td>
<td>2.535</td>
<td>5.375</td>
<td>33</td>
<td>5.305</td>
<td>3.611</td>
<td>2.722</td>
</tr>
<tr>
<td>34</td>
<td>1.775</td>
<td>4.767</td>
<td>4.637</td>
<td>4.954</td>
<td>35</td>
<td>4.610</td>
<td>4.988</td>
<td>2.760</td>
</tr>
<tr>
<td>36</td>
<td>4.922</td>
<td>2.555</td>
<td>1.069</td>
<td>0.136</td>
<td>37</td>
<td>4.872</td>
<td>2.380</td>
<td>3.450</td>
</tr>
<tr>
<td>40</td>
<td>1.876</td>
<td>2.829</td>
<td>4.350</td>
<td>4.341</td>
<td>41</td>
<td>2.862</td>
<td>3.822</td>
<td>4.120</td>
</tr>
<tr>
<td>44</td>
<td>3.634</td>
<td>4.737</td>
<td>1.241</td>
<td>1.659</td>
<td>45</td>
<td>2.793</td>
<td>3.333</td>
<td>4.482</td>
</tr>
<tr>
<td>48</td>
<td>3.555</td>
<td>1.838</td>
<td>1.613</td>
<td>0.228</td>
<td>49</td>
<td>2.935</td>
<td>1.992</td>
<td>3.177</td>
</tr>
<tr>
<td>50</td>
<td>1.519</td>
<td>2.472</td>
<td>1.263</td>
<td>3.579</td>
<td>51</td>
<td>3.709</td>
<td>1.302</td>
<td>4.083</td>
</tr>
<tr>
<td>52</td>
<td>2.739</td>
<td>3.301</td>
<td>3.059</td>
<td>4.041</td>
<td>53</td>
<td>6.412</td>
<td>1.812</td>
<td>2.736</td>
</tr>
<tr>
<td>54</td>
<td>2.241</td>
<td>2.399</td>
<td>2.290</td>
<td>3.138</td>
<td>55</td>
<td>1.632</td>
<td>1.361</td>
<td>2.906</td>
</tr>
<tr>
<td>58</td>
<td>2.912</td>
<td>3.803</td>
<td>2.920</td>
<td>3.928</td>
<td>59</td>
<td>4.461</td>
<td>2.121</td>
<td>2.607</td>
</tr>
<tr>
<td>60</td>
<td>4.166</td>
<td>4.776</td>
<td>1.570</td>
<td>1.602</td>
<td>61</td>
<td>4.369</td>
<td>1.913</td>
<td>5.348</td>
</tr>
<tr>
<td>62</td>
<td>3.701</td>
<td>2.772</td>
<td>2.023</td>
<td>2.060</td>
<td>63</td>
<td>3.998</td>
<td>4.294</td>
<td>1.005</td>
</tr>
<tr>
<td>64</td>
<td>3.833</td>
<td>1.691</td>
<td>4.080</td>
<td>0.353</td>
<td>65</td>
<td>4.793</td>
<td>3.640</td>
<td>3.629</td>
</tr>
<tr>
<td>66</td>
<td>3.761</td>
<td>2.415</td>
<td>3.361</td>
<td>1.377</td>
<td>67</td>
<td>2.196</td>
<td>2.529</td>
<td>3.630</td>
</tr>
<tr>
<td>68</td>
<td>1.831</td>
<td>4.207</td>
<td>4.744</td>
<td>4.914</td>
<td>69</td>
<td>2.406</td>
<td>4.228</td>
<td>1.710</td>
</tr>
<tr>
<td>70</td>
<td>3.771</td>
<td>1.848</td>
<td>4.274</td>
<td>0.010</td>
<td>71</td>
<td>1.295</td>
<td>4.452</td>
<td>4.500</td>
</tr>
<tr>
<td>72</td>
<td>1.214</td>
<td>2.058</td>
<td>2.738</td>
<td>3.874</td>
<td>73</td>
<td>3.442</td>
<td>1.937</td>
<td>4.913</td>
</tr>
<tr>
<td>74</td>
<td>3.073</td>
<td>4.637</td>
<td>1.040</td>
<td>1.699</td>
<td>75</td>
<td>3.907</td>
<td>2.881</td>
<td>1.974</td>
</tr>
</tbody>
</table>

\( t = t_{15}; W(t_1) - W(t_2) = 6.198 - 0.001 < 19.38 \Rightarrow t = t_{15}; W(T) = 0.001 \)