J Popul Econ (1994) 7:333–350

Journal of
Population Economics

The measurement of household cost functions
Revealed preference versus subjective measures

Arie Kapteyn *

Center for Economic Research, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands

Received March 7, 1994 / Accepted August 10, 1994

Abstract. Since the work of Pollak and Wales (1979), it is well-known that demand data are insufficient to identify a household cost function. Hence additional information is required. For that purpose I propose to employ direct measurement of feelings of well-being, elicited in surveys.

In the paper I formally establish the connection between subjective measures and the cost function underlying the AID system. The subjective measures fully identify cost functions and the expenditure data do this partly. This makes it possible to test the null hypothesis that both types of data are consistent with one another, i.e. that they measure the same thing. I use two separate data sets to set up a test of this equivalence. The outcomes are somewhat mixed and indicate the need for further specification search. Finally, I discuss some implications of the outcomes.

1 Introduction

Household cost functions (and equivalence scales) can have many purposes and many underlying assumptions, as for instance stressed by Browning (1993). (See also Nelson (1993) for a historical and philosophical account.) In this paper I am concerned with the question how household cost functions depend on the composition of a household. Traditionally, a question like this is answered by the in-

* Presidential address delivered at the seventh annual meeting of the European Society for Population Economics, June 2–5, Budapest, Hungary. A first version was presented at the meetings of the Allied Social Sciences Associations, Anaheim, California, January 5–7, 1993. The author thanks Rob Alessie for help with the data, and the Netherlands Central Bureau of Statistics for their permission to use the data. Both Rob Alessie and an anonymous referee made very helpful comments on an earlier version.

corporation of demographic factors in demand systems. As has been argued by Pollak and Wales (1979) one cannot fully identify household cost functions from demand data alone. Although this is not a problem in all cases, e.g. if one only wants to use a household cost function as a representation of preferences from which to derive demand equations, it does pose problems if one wants to use cost functions in applied welfare analysis.

The most obvious solution to an identification problem is to invoke additional information. It can be argued that, rather than employing data on consumption expenditures, a household's cost function can also be measured, and with less effort, by asking respondents to a survey subjective questions about money amounts needed to attain a certain welfare level. This approach has been adopted by a limited number of authors including Kapteyn and Van Praag (1976), Kapteyn et al. (1988), Hagenaars (1986), Van Praag and Van der Sar (1988), Dubnoff (1979), Vaughan (1984), Danziger et al. (1984), Colasanto et al. (1984), De Vos and Garner (1991). Although in my opinion this *direct measurement* has proven to work very well and to yield sensible results, it is fair to say that the profession of economists has generally ignored the direct approach.

I am not entirely sure why this is. On the basis of my own discussions with other economists (including discussants at conferences and referees for journals) I would conjecture that most economists simply do not believe what people say. They feel that the questions asked to respondents are too difficult or abstract to yield sensible answers. Hence they cannot believe that what people say reflects preferences in the same way that observed choice behavior does. And if responses to questions do not measure the same thing as observed behavior then direct measurement becomes *irrelevant* for empirical economics. This impression is probably reinforced by the feeling that direct measurement yields results that appear different than outcomes obtained through demand analysis (which I will henceforth refer to as *indirect measurement* or the *revealed preference approach*).

The purpose of this paper is to formally test whether direct and indirect measurement of cost functions are equivalent, i.e. whether the two approaches measure the same concept. This is important for various reasons. In the first place direct measurement is much simpler than indirect measurement and hence more cost effective. So if we can accept the hypothesis that both modes of measurement measure the same thing, empirical analysis may be greatly facilitated. Secondly, the direct approach does not suffer from the same identification problem as the indirect approach. Hence if we can accept equivalence, we also solve a fundamental problem that has been bugging applied welfare analysis. In the third place, as will become clear below, combination of indirect and direct approaches yields new possibilities for the detection of misspecification in empirical models and solution of the ensuing problems.

For a start, I will present an example in Sect. 2 illustrating the identification problem inherent in the revealed preference approach. In Sect. 3 I will provide a brief discussion of the informational requirements for identification. In Sect. 4 I will introduce the cost function of the Almost Ideal Demand System, which will serve as the main vehicle for setting up an empirical test. There I also discuss indirect measurement. In Sect. 5 direct measurement of the same cost function is described. In Sects. 6 and 7 I develop the (simple) econometric framework that will allow for a test of equivalence of direct and indirect measurement. In Sect. 8 the outcomes of the test are presented. A discussion of the results and their implications follows in Sect. 9.

2 Underidentification of cost functions; an example

Consider the following two utility functions:

$$U(q,f) \;=\prod_{i=1}^{k}(q_i-a_i)^{\beta_i} \tag{2.1}$$

$$U^*(q,f)=\sum_{i=1}^{k}\beta_i\ln(q_i-a_i)+\varepsilon' f \tag{2.2}$$

where q: = k-vector of goods, f: = vector of household characteristics, a_i, β_i: = parameters, *which may depend on* f, ε: = parameter vector.

Maximization of either of these functions with respect to q, subject to a linear budget constraint yields the following demand functions:

$$p_i q_i = p_i a_i + \beta_i\left(x-\sum_{j=1}^{k}p_j a_j\right), \quad i=1,\ldots,k \tag{2.3}$$

where p_i: = prices, $i=1,\ldots,k$; x: = total expenditures.

The reason why the utility functions U and U^* yield the same demand functions is obvious. U^* is equal to the log of U plus a constant $\beta' f$. Hence, if U reaches a maximum, so does U^*.

By substituting the demand equations into the utility function we can easily derive the cost functions associated with U and U^*. They are, respectively:

$$c(u,p,f) \;= u\cdot\prod_{i=1}^{k}\left(\frac{p_i}{\beta_i}\right)^{\beta_i}+\sum_{i=1}^{k}p_i a_i \tag{2.4}$$

$$c^*(u^*,p,f)=e^{u^*}\cdot\prod_{i=1}^{k}\left(\frac{p_i}{\beta_i}\right)^{\beta_i}\cdot e^{-\varepsilon' f}+\sum_{i=1}^{k}p_i a_i \tag{2.5}$$

If demand data are available one can estimate all parameters in the demand equation (2.3). If these parameters depend on household characteristics then the parameters appearing in the relation between the demand parameters and household characteristics can be estimated as well. As indicated above, I will refer to this way of measurement of cost function parameters as *indirect measurement* or *revealed preference measurement*. Clearly, the parameter vector ε in (2.5) cannot be identified from the demand equation, because ε does not appear in the demand equation. Nor is it possible to tell wether c or c^* is the correct cost function.

Although I have chosen to illustrate the identification problem by means of an example, it should be clear that the problem is perfectly general. Demand data alone can never identify a household cost function completely.

3 Informational requirements

The fact that demand data are not sufficient to identify a cost function completely was first noted by Pollak and Wales (1979), and later reiterated by Lewbel (1989), Fisher (1987), Blackorby and Donaldson (1988), Pashardes (1992), and others. Whenever one faces an identification problem, there are three basic choices. The first is to accept the problem and to try and live with it. This includes an attempt to see what can still be salvaged from the wreckage. The second approach is to make arbitrary assumptions that (seemingly) make the problem go away. The third approach is to invoke additional information. I will briefly discuss these three approaches in the present context, borrowing freely from Blundell and Lewbel (1991).

- *Trying to live with it.* Blundell and Lewbel (1991) prove a beautiful lemma which says that within a given price regime *any* equivalence scale (i.e. the cost of living of one household relative to another) is consistent with observed demand. That is, equivalence scales are not identified. At the same time the evolution of these equivalence scales with changes in the price regime is fully identified. One can paraphrase this by saying that we can fully identify the changes in something that we cannot see. I doubt if there are many contexts in which such information is useful. A referee makes the following comment about this: "I am more positive about the Blundell-Lewbel result on updating equivalence scales than the author. It seems to me that the result is useful once we agree on a scale in the base year. Essentially it formalises the obvious point that whatever the scale, it should increase if the price of milk increases." Of course this is true, but it appears to me that the main problem remains to agree on a scale in the base year.
- *Arbitrary assumptions.* If no extra information is invoked (see below), any assumption that solves the identification problem is by definition arbitrary. Many assumptions have been made in the literature either implicitly or explicitly. A popular assumption has been the *Independence of Base* (IB) assumption (or equivalence scale exactness, as it is denoted by Blackorby and Donaldson 1988), which stipulates that the ratio of cost functions for two households is independent of the level of utility at which the cost functions are evaluated. Although IB places testable restrictions on observable demands, acceptance of these restrictions does not solve the identification problem completely. This can be illustrated by the L. E. S. example in the previous section. IB implies for both (2.4) and (2.5) that the parameters a_i have to be equal to zero. One can see immediately from (2.3) that this is a testable hypothesis. However, if this hypothesis is accepted by the data, and if we are therefore willing to maintain that the parameters a_i are zero, this does not imply that equivalence scales can be identified. There is still no way to choose between (2.4) and (2.5). We have to make the additional, untestable, assumption that all monotonic transformations of u that are allowed in (2.4) do not involve household composition. In other words, in (2.5) the vector ε has to be identically equal to zero. Clearly in that case (2.4) and (2.5) will yield identical equivalence scales.

  So, acceptance of IB does not solve our problems. On the other hand, if IB is rejected then even the additional assumption that u is uniquely determined up to monotonic transformations not involving household composition, does not determine equivalence scales uniquely. It is worthwhile therefore to note that tests of IB by Blundell and Lewbel (1991) and by Pashardes (1992) indicate sound rejection.

One can also formulate IB, or exactness, in terms of *differences* of cost functions rather than ratios (Blackorby and Donaldson 1993). In that case the difference between cost functions of different households should not depend on utility. In the L. E. S. example one can see that IB in this sense will hold for (2.4) if the β_i do not depend on household composition. The difference in cost of two households h and r say is then simply:

$$\sum_{i=1}^{k} p_i(a_{ih}-a_{ir}) \ , \tag{3.1}$$

where the subscripts h and r indicate dependence of the parameters on the composition of the households h and r. This outcome remains unaffected if we allow transformations of u not involving household composition. In other words in (2.5) the vector ε has to be identically equal to zero. I ignore the pathological case that all β_i are zero. In that case the utility function is a constant.

The IB assumption is by no means the only assumption that can be made to avoid the identification problem. But all assumptions have in common that they are arbitrary if we do not invoke additional information.

- *Additional information.* Blundell and Lewbel mention two types of additional information one could conceive of. The first type is to have observations on revealed preference for household compositions. Although one can conceive of such an approach in principle, it certainly stretches one's imagination as to how this would have to be implemented in practice. The other possibility they mention is the use of direct questions on household cost functions. And that is the approach I want to pursue in the rest of this paper.

4 The cost function of the Almost Ideal Demand system

For concreteness the rest of the analysis will be done for a specific choice of functional form. Consider the formula for a PIGLOG cost function (Muellbauer 1975):

$$\ln(c(u,\boldsymbol{p})) = a(\boldsymbol{p})+b(\boldsymbol{p})u \ , \tag{4.1}$$

where $\boldsymbol{p}$ is a vector of prices and $a(\boldsymbol{p})$ and $b(\boldsymbol{p})$ are functions of prices. Furthermore, let us specialize the PIGLOG formulation to the Almost Ideal Demand specification of Deaton and Muellbauer (1980) and define

$$a(p) = \alpha_0 + \sum_k \alpha_k \ln(p_k) + \frac{1}{2}\sum_k \sum_l \gamma_{kl} \ln(p_k)\ln(p_l) \tag{4.2}$$

$$b(p) = \beta_0 \prod_k p_k^{\beta_k} \ , \tag{4.3}$$

where the parameters α_k, γ_{kl}, β_k have to satisfy well-known homogeneity restrictions. This gives rise to demand equations of the form:

$$w_i = \alpha_i + \sum_j \gamma_{ij} \ln(p_j) + \beta_i [\ln(x) - a(p)] \; , \tag{4.4}$$

where w_i is the budget share of the i-th commodity, $i = 1, \ldots, I$; x is total expenditures. The parameters in the demand system, and hence the parameters of the cost function, can be estimated if one has data available on the consumption of households under a sufficiently rich variation in prices. That is, in this way the parameters are measured indirectly, as defined in Sect. 2.

Clearly, for the cost function (4.1) to satisfy IB in a relative form the parameters in $b(p)$ should not depend on household composition. This is a testable proposition. If IB is satisfied, equivalence scales would be identified from demand data if furthermore u would be determined up to monotonic transformations not depending on household composition. And, as with the L.E.S. example, there is no way of knowing whether this is true without additional information. It is to such additional information that I now turn.

5 Direct measurement

In the literature around the so-called individual welfare function of income (WFI), spawned by Van Praag (1968), much of the empirical analysis is based on the answers to the following question:

Which after tax income would you in your circumstances consider to be very bad? And bad? Insufficient? Sufficient? Good? Very good? (We mean after tax *household* income)

very bad $ ——
bad $ ——
insufficient $ ——
sufficient $ ——
good $ ——
very good $ ——

If one accepts the verbal qualifications "good", "sufficient", "bad", etc. as indications of utility levels the IEQ measures a cost function *directly*. For, the answers then provide for each of the utility levels the amount of money required to attain that utility level. Since the preamble states that answers have to be given "in your circumstances" the cost function is measured conditional on these circumstances. There is of course some ambiguity as to what these circumstances are, but one would expect family composition to be one of them.

In the WFI-literature a specific functional form for the cost function is assumed, corresponding to an indirect utility function that has a lognormal shape $\Lambda(.;\mu,\sigma)$. To measure the parameters μ and σ of the lognormal utility function for a given respondent it is commonly assumed that the verbal qualifications in the IEQ can be transformed into numbers, say $e_i, i = 1, \ldots, 6$, between zero and one. These numbers partition the [0,1] interval in equal intervals, i.e. $e_i = (2i-1)/12$. In other words, the label "very bad" is associated with $e_1 = 1/12$, the label "bad" with $e_2 = 3/12$, etc. If we denote the answers given by a respondent by z_i, $i = 1, \ldots 6$, then by assumption the answers satisfy approximately

$$N(\ln(z_i); \mu, \sigma) = N\left(\frac{\ln(z_i) - \mu}{\sigma}; 0, 1\right) = e_i, i = 1, \ldots, 6 \ . \tag{5.1}$$

This implies that approximately,

$$\frac{\ln(z_i) - \mu}{\sigma} = N^{-1}(e_i; 0, 1) \ , \quad i = 1, \ldots, 6 \ . \tag{5.2}$$

Adding an error term to allow for measurement and rounding errors in the answers of a respondent, the parameters μ and σ of an individual can now be estimated by the following regression:

$$\ln(z_i) = \mu + \sigma N^{-1}(e_i; 0, 1) + \varepsilon_i \ . \tag{5.3}$$

Further details are for instance given in Van Praag (1971) and Van Praag and Kapteyn (1973). Since this mode of measurement was introduced, various tests of the underlying assumptions have been carried out, including the equal interval assumption and lognormality (see, e.g., Antonides et al. 1980; Van Herwaarden and Kapteyn 1981; Buyze 1982; Van Praag 1991). The outcomes of these tests are not uniformly supportive of the underlying assumptions, but they indicate their approximate validity.

Since by assumption $N^{-1}(e_i; 0, 1)$ is nothing else than a positive monotonic transformation of a utility level, and since there is no presumption that μ and σ do not depend on prices, we might as well write (5.3) as

$$\ln(z_i) = \mu(p) + \sigma(p) u_i + \varepsilon_i \ , \tag{5.4}$$

where $u_i = N^{-1}(e_i; 0, 1)$. Comparing this to the PIGLOG cost function given by (4.1), suggests that the IEQ may be seen to measure a PIGLOG cost function by means of direct questions rather than through observation of behavior.

Similarly the analogy of (4.1) and (5.4) suggests that

$$\mu(p) = \alpha_0 + \sum_k \alpha_k \ln(p_k) + \frac{1}{2} \sum_k \sum_l \gamma_{kl} \ln(p_k) \ln(p_l) \tag{5.5}$$

$$\sigma(p) = \beta_0 \prod_k p_k^{\beta_k} \ . \tag{5.6}$$

Since μ and σ can be measured per individual, one could estimate the parameters on the right hand side of (5.5) and (5.6) by regressing μ and σ on the functions of prices on the right hand side of (5.5) and (5.6). So this then amounts to the *direct measurement* of the parameters of the AID cost function.

So we now have two ways to measure the parameters of the AIDS, namely through the observation of demand (i.e. through revealed preference) or through direct measurement. It is this fact that allows us to test in principle whether the direct measurement and the revealed preference approach measure the same thing.

6 Econometric implications

There are at least two reasons why testing for the equivalence of the direct and the indirect approach is less straightforward than a comparison of (4.4) and (5.5)–(5.6) would suggest. The first reason is that the models are not complete; most likely preferences vary across households. In the present set-up I ignore the possibility that households are not homogeneous decision making units. Thus I assume that both observed consumption behavior and answers to the IEQ either reflect household preferences or the preferences of the dictator in the household. Neglect of such variation may bias the test. A second reason is that no data sets exist that permit both the estimation of the demand system and the measurement of WFIs.

Turning to the first problem, I assume that the following simple equation provides a sufficiently accurate description of the variation of preferences across households.

$$\alpha_{0,n} = \delta_0 + \delta' f_n + \xi_n \ , \tag{6.1}$$

where n now indexes the household, $\alpha_{0,n}$ is simply α_0 as occurring in (4.2), but with an index n to indicate that it may vary across households. δ_0 and δ are parameters, f_n is a vector of household characteristics for household n; ξ_n represents all other factors that may influence the household's preferences. These factors may include reference group effects, habit formation, random effects, etc. In itself this modelling of preference variation may be far too restrictive to be adequate. In the concluding section I return to this issue.

Let us rewrite (4.4) by indexing all variables by n and adding an i.i.d. error term u_{ni}, and by replacing $a(p)$ by $a_n(p)$, where $a_n(p)$ is defined according to (4.2), but with α_0 replaced by $\alpha_{0,n}$. So we obtain:

$$w_{ni} = \alpha_i + \sum_j \gamma_{ij} \ln(p_{nj}) + \beta_i [\ln(x_n) - a_n(p)] + u_{ni} \ . \tag{6.2}$$

Similarly, we replace (5.5) by

$$\mu_n = \alpha_{0,n} + \sum_k \alpha_k \ln(p_{nk}) + \frac{1}{2} \sum_k \sum_l \gamma_{kl} \ln(p_{nk}) \ln(p_{nl}) + v_n = a_n(p) + v_n \tag{6.3}$$

with v_n an error term, representing for instance measurement error in μ_n.

It is worth commenting on the effect of the variable ξ_n implicit in $a_n(p)$. First of all we should note that if ξ_n is not fully specified this introduces bias in the estimates of the parameters of the two equations above, unless we could claim that the omitted factors do not correlate with the included explanatory variables. In general the bias will be different in the two equations and hence the direct and the indirect approach to measurement of cost functions yield different outcomes. The only way to avoid the omitted variable bias is to have a complete specification of all factors influencing preferences. With respect to the explanation of variation in the welfare parameters μ and σ across households numerous papers have been written documenting these various influences (see, e.g., Kapteyn et al. 1985 or Hagenaars 1986). The literature on taste shifting in demand systems is relatively less voluminous, but also here the evidence points at significant effects. See, e.g., Alessie and Kapteyn (1991) for evidence that in the AID system demographic effects, habit formation, and reference group effects all play a role. This evidence also suggests that ξ will be correlated with most if not all explanatory variables in a demand system like (4.4). This strengthens the observation that omission of relevant factors will bias estimates.

So how can we devise a test of the equivalence of direct and indirect measurement that is not affected by this omitted variable bias? Note that under the adopted formulation the null hypothesis of equivalence of direct and indirect measurement implies that

$$w_{ni}+\beta_i\mu_n = \alpha_i + \sum_j \gamma_{ij}\ln(p_{nj}) + \beta_i \ln(x_n) + u_{ni} + \beta_i v_n \ . \tag{6.4}$$

A test of the null can now take the form of adding f_n and bilinear functions of log-prices to the right hand side of (6.4) and testing for significance of their coefficients. For later treatment it is useful to consider a particular alternative hypothesis, namely that the equation for μ_n reads

$$\mu_n = \alpha^*_{0,n} + \sum_k \alpha^*_k \ln(p_{nk}) + \frac{1}{2}\sum_k \sum_l \gamma^*_{kl}\ln(p_{nk})\ln(p_{nl}) + v_n \tag{6.5}$$

and

$$\alpha^*_{0,n} = \delta^*_0 + \delta^{*\prime} f_n + \xi_n \ . \tag{6.6}$$

In other words, the functional form is the same as under the null, but the parameters are different. This leads to

$$\begin{aligned} w_{ni}+\beta_i\mu_n = {} & \alpha_i + \beta_i(\delta^*_0-\delta_0) + \sum_j [\gamma_{ij}+\beta_i(\alpha^*_j-\alpha_j)]\ln(p_{nj}) + \beta_i\ln(x_n) \\ & + \beta_i(\delta^*-\delta)' f_n + \frac{1}{2}\beta_i \sum_k\sum_l (\gamma^*_{kl}-\gamma_{kl})\ln(p_{nk})\ln(p_{nl}) \\ & + u_{ni} + \beta_i v_n \ . \end{aligned} \tag{6.7}$$

In obvious notation this can be written with "reduced form coefficients" as

$$\begin{aligned} w_{ni}+\pi_{4,i}\mu_n = {} & \pi_{0,i} + \sum_j \pi_{1,ij}\ln(p_{nj}) + \pi'_2 f_n + \sum_k\sum_l \pi_{3,kl}\ln(p_{nk})\ln(p_{nl}) \\ & + \pi_{4,i}\ln(x_n) + u_{ni} + \pi_{4,i} v_n \ . \end{aligned} \tag{6.8}$$

Under the null, we have that $\pi_2 = 0$ and $\pi_{3,kl} = 0$. So if data on all variables in (6.8) were available, we could simply run a regression and apply F- or t-tests to test the null. Since, as mentioned above, no single data set is available containing all variables in (6.8) we have to combine different samples.

7 Combining samples

Two datasets are available. The first one is a consumer expenditure panel which has run from April 1984 through September 1987. (This is the so-called Intomart consumer index; the data used here were prepared by Pim Adang.) This panel allows for the estimation of a demand system, including demographics, using monthly observations, but does not allow for the measurement of WFIs. The second dataset is a household panel measuring income, labor market status, demographics, and the like. (This is the so-called socio-economic panel run by the Netherlands Central Bureau of Statistics. In this paper I use an extract from the data constructed by Alessie et al. 1992.) Also, WFIs are measured. The interviews have taken place in October 1984, October 1985 and October 1986. I will refer to the first panel as the CEP (consumer expenditure panel) and to the second panel as the SEP (socio-economic panel).

In the empirical work I shall consider only two goods, "food" and "other". Monthly price indices can be constructed from official statistics. In view of the fact that only two commodities are considered and given the homogeneity restrictions on coefficients in the AID system, only the relative price index of "food" relative to "other" enters the demand system. Also, we only have to consider one equation from the system, as the other follows from adding up. This allows us to drop the subscript i and to write (6.8) as

$$w_n+\pi_4\mu_n=\pi_0+\pi_1\ln(p_n)+\pi_2' f_n+\pi_3[\ln(p_n)]^2+\pi_4\ln(x_n)+u_n+\pi_4 v_n \ . \tag{7.1}$$

In the estimation of Eq. (7.1) I follow the recent literature on the combination of samples (see, e.g., Arellano and Meghir 1991; Angrist and Krueger 1992; Lusardi 1993). Simplify Eq. (7.1) even further by writing it in matrix format as

$$w+\mu\pi_4=X_1\theta+X_2\pi_4+\varepsilon \ , \tag{7.2}$$

where X_1 is a matrix containing a column of ones plus the observations on the first three variables on the right hand side of (7.1) and X_2 is a vector containing the observations on log-expenditures. The parameter vector θ is defined as $\theta=(\pi_0,\pi_1,\pi_2',\pi_3)'$.

If all variables were observed for all households, and if Z were a matrix of valid instruments one would typically estimate the parameters of interest by constructing the vector

$$Z'(w+\mu\pi_4-X_1\theta-X_2\pi_4) \tag{7.3}$$

and minimizing its length with respect to π_4 and θ in some appropriate metric. Note that total expenditures are probably not statistically exogenous, and hence instrumental variable estimation is required.

The elements of X_1 are observed for both samples, but the elements of X_2 and w are only observed for the CEP, whereas the elements of μ are only observed for the SEP. Let Z_c be a matrix of instruments observed for the CEP sample and let Z_s contain observations on the same instruments for the SEP sample. If both samples can be considered to be drawings from the same population then consistent estimation of the parameters can take place by minimizing the length of the following vector:

$$\frac{1}{N_c} Z_c' w + \frac{1}{N_s} Z_s' \mu \pi_4 - \frac{1}{N} Z' X_1 \theta - \frac{1}{N_c} Z_c' X_{2c} \pi_4 \ , \tag{7.4}$$

where Z without subscript and X_1 stands for the matrix of instruments and variables for both samples combined. N_c is the number of observations in the CEP and N_s is the number of observations in the SEP, $N = N_c + N_s$. The minimization problem can be solved in a very simple way. This can be seen as follows. Define the vector $y = \left(\frac{N}{N_c} X_{2c}', -\frac{N}{N_s}\mu\right)'$ and the vector $z = \left(\frac{N}{N_c} w', 0\right)'$. Then we can rewrite the above vector as follows:

$$\frac{1}{N} Z'(z - X_1 \theta - y \pi_4) \ , \tag{7.5}$$

which we recognize as the vector that would be minimized if we would estimate the following model by instrumental variables:

$$\frac{1}{N} z = \frac{1}{N} X_1 \theta + \frac{1}{N} y \pi_4 + \text{error} \ . \tag{7.6}$$

The only thing that remains to be done for efficient estimation is to derive the asymptotic variance covariance matrix of the *error*. This is done in the Appendix. With this covariance matrix in hand one can apply generalized least squares.

8 Empirical results

For 91 households in the CEP observations are available for all 42 months that the panel has been in existence. Thus we have 3822 observations in total. The balanced panel extracted from the SEP has 1328 households. This number is much lower than would be possible, since the SEP covers approximately 5000 households. However in Alessie et al. (1992) a severe selection has been made, since extensive information on households' reference groups had to be available. For simplicity I have not tried to construct a bigger sample. Since three waves are used in the empirical analysis, we have 3984 observations. At first sight, issues of selectivity and individual and time effects would appear to complicate the analysis. However, to the extent that these effects would only affect the distribution of ξ_n in (6.1) the set-up of model (6.4) essentially wipes out all such effects. I return to this in the next section. Hence we use the observations as if they are independent, conditional on the exogenous variables in the model.

Only a limited number of variables can be used as instruments, since the definition of variables across samples appears to differ widely. It turns out that only degree of urbanization and province of residence are defined in identical ways for the two samples. For the rest I consider prices and household composition as exogenous, so these yield valid instruments as well. The influence of household composition has been modelled in an extremely simple way, namely as the log of the number of family members. This may appear too primitive, but it does not bias the test under the null. For, any misspecification in the modelling of the influence of family composition will be absorbed by the variable ξ_n in (6.1), which does not appear in the test.

In Table 1 I present four sets of results. In the first column the results of estimating a food share equation analogous to (6.2) are presented. In the second column estimates are given obtained by estimating Eq. (7.6) by OLS. In the third and fourth column I present the estimates obtained by the IV approach outlined above. The difference between the latter two columns lies in the definition of instruments. In the third column urbanization degree has been defined as a set of six binary variables with province a variable with domain 1, . . . , 11. In the fourth column urbanization degree has been defined identically, but province has now been defined as a set of eleven binary variables. Although both definitions of instruments would appear to be valid choices, one would expect the latter set of instruments to be superior in terms of the asymptotic efficiency of the resulting estimators.

Recently a number of authors have pointed at the danger of using weakly correlated instruments, because the usual asymptotic theory can be severely misleading (e.g. Bound et al. 1993; Staiger and Stock 1993; Bekker 1994). One way of looking at this is to consider the R^2 of the regression of the potentially endogenous variable on the instruments (in our case y in Eq. (7.6), i.e. the vector containing the values of log-expenditures and μ). For column 3 this R^2 equals 0.10, whereas for column 4 we obtain 0.13. These values seem to be sufficiently high to allow for the application of standard asymptotic theory.

In view of the purpose of this paper, the most striking aspect of Table 1 is that the variable ln (fs), which is highly significant in the food share equation and also comes out highly significant when estimating (7.6) with OLS becomes totally insignificant when estimating the model by means of IV, as in the third column.

Table 1. Estimates for three specifications

Variable	Food share	Two samples OLS	Two samples IV1	Two samples IV2
ln (p)	−0.829	−1.94	−4.91	−2.51
s.e.	0.390	0.508	0.815	0.694
t	−2.12	−3.88	−6.03	−3.62
ln (fs)	0.069	0.031	0.003	0.026
s.e.	0.006	0.004	0.007	0.005
t	11.1	7.710	0.359	5.73
$\ln^2 (p)$	−20.3	−51.2	−120	−62.9
s.e.	7.83	11.9	19.0	17.4
t	−2.59	−4.35	−6.32	−3.62
ln $(x)/\mu$	−0.084	0.010	0.007	0.010
s.e.	0.013	0.00008	0.0006	0.006
t	−6.36	120	12.5	15.8
Constant	1.22	0.22	0.228	0.215
s.e.	0.159	0.005	0.006	0.004
t	7.73	40.7	37.5	48.1
R^2	0.36	0.665		
# of obs.	3822	7806	7804	7804

In the fourth column however, with the use of the more efficient instruments, the coefficient of $\ln(fs)$ is once again highly significant though smaller in absolute value than with OLS.

It should be noted furthermore that the variable $\ln^2(p)$ remains significant in both IV-columns, whereas according to (6.4) this variable should become insignificant as well. Although this is at variance with the null as formulated so far, it is easy to think of a cost function which would be compatible with a significant $\ln^2(p)$ variable. That would still be a PIGLOG cost function but with a function $a(p)$ defined as a cubic polynomial in log-prices rather than as a quadratic (cf. (4.2)):

$$a(p) = \alpha_0 + \sum_k \alpha_k \ln(p_k) + \frac{1}{2} \sum_k \sum_l \gamma_{kl} \ln(p_k) \ln(p_l)$$
$$+ \frac{1}{3} \sum_k \sum_l \sum_m \varphi_{klm} \ln(p_k) \ln(p_l) \ln(p_m) \ . \tag{8.1}$$

This cost function would also imply the presence of $\ln^3(p)$ in the foodshare equation. I have estimated the foodshare equation as in column one of Table 1, but with $\ln^3(p)$ included. It turns out that the fit of the equation does not change. The reason for this is simply that in the present data set the variables $\ln(p)$, $\ln^2(p)$ and $\ln^3(p)$ are highly collinear: A regression of $\ln^3(p)$ on $\ln(p)$ and $\ln^2(p)$ yields an R^2 equal to 0.994. In fact the fit of the foodshare equation is identical whether we include $\ln(p)$ and $\ln^2(p)$ or $\ln(p)$ and $\ln^3(p)$.

I have also performed a test of overidentification of the instruments by regressing the residuals of the estimated equations according to columns 3 and 4 on the instruments used. We can use N times the R^2 of this regression as a test statistic for the null that the instruments province and urbanization degree should not enter (7.6) directly. The test statistic should follow a χ^2 distribution with $l-k$ degrees of freedom where l is the number of instruments and k the number of explanatory variables in the model (see e.g. Davidson and MacKinnon 1993). For both columns 3 and 4 the null is soundly rejected ($\chi^2(10) = 30.4$ and $\chi^2(20) = 86.8$ respectively). Next I have experimented with the inclusion of some instruments directly. Although, of course this does reduce the value of the χ^2 statistic the values of the coefficients in Table 1 remain essentially unchanged. Another point worth noting is that the estimate of β (the coefficient of $\ln(x)$ and μ respectively) is positive in all columns but the first. Since food is generally considered to be a necessity β should come out negative. This presents another anomaly that points at misspecification.

In sum, we find that we can formulate a specification of the cost function such that according to one set of instruments the null hypothesis that direct and indirect measurement are equivalent would pass the test, whereas a different (more efficient) set of instruments yields the conclusion that the two modes of measurement are not equivalent, although the size of the coefficient of log-family size in column four suggests that the null hypothesis may be reasonably close to the truth. At the same time there are indications of misspecification in all cases.

9 Discussion

The empirical analysis has been based on a rather simple model. This leads to two sorts of considerations. First of all, under the null, misspecification due to an overly simplistic set-up, e.g. the representation of family composition merely by the log of family size, is absorbed by the variable ξ_n in (6.1) and hence should not bias the test of the null.

A second kind of consideration is that if the model chosen is too simplistic, then this misspecification will tend to be picked up by variables added to the model, even if these variables do not properly belong to the model. In other words one tends to obtain too many significant variables. Since my test is based on precisely the addition of variables to an equation that, under the null should not be there, the test would seem to be biased against the null. An example of a likely source of misspecification is the disregard of issues of selectivity and serial correlation; to the extent that these enter the equations through the error term in (6.1), they are wiped out by the combination of (6.2) and (6.3) into (6.4). To the extent that selectivity and serial correlation affect the equations in a different way, one would expect the model (6.4) to be misspecified. This misspecification may then be picked up by the variable $\ln(fs)$, and hence the test will be biased against the null. This would then explain the significant coefficient in the fourth column of Table 1. Yet another source of misspecification would occur if the functional form of the share equation for food were inadequate. Evidence provided by Banks et al. (1994) suggests that the AID share equation is sufficiently general to adequately describe the demand for food.

Altogether then the evidence appears to be a bit mixed. Formally, the null is rejected, but the estimated coefficient of $\ln(fs)$ is not very large. Given the various sources of misspecification mentioned this is about what one would expect if the null were true. Hence, although the issue of equivalence of direct and indirect measurement of cost functions has not been settled by the simple test I have proposed here, further research into the hypothesis seems justified. Among other things, one may consider more complex specifications than (6.1).

An important aspect of the test applied here is that it tries to deal with omitted variables. It is readily seen that omitted variables lead to different biases in a demand equation than in for instance (6.3). As noted in Sect. 6, this implies that equivalence scales derived from demand systems will be different from scales derived from subjective measures. These differences may simply point to misspecification rather than to genuine differences between direct and indirect measurement. This is not to say that demand systems and subjective measures will only suffer from similar sources of misspecification. In Kapteyn et al. (1988) specific methodological issues in the application of subjective measures are being discussed. Their correction method has been used to construct the values for μ in the current data set.

One should also note that equivalence scales show enormous variation across studies based solely on demand data (Browning 1992). This variation itself may point to misspecification in the models considered, if only because the models cannot all be true at the same time. More importantly, since all equivalence scales based on demand data suffer from the identification problem alluded to in Sect. 2, one may claim that the scales obtained by various authors are inherently arbitrary. Recall the Lemma proven by Blundell and Lewbel (1991), quoted in Sect. 3 above.

Imagine that the null hypothesis put forward in this paper were accepted as being true, then this would have a number of consequences. First of all it would suggest that the particular representation of the utility function adopted here is adequate. Hence, if one were able to fully specify the AID system (with third order terms in log-prices, and not omitting relevant variables) equivalence scales could be estimated that are not arbitrary. Secondly, however, the outcomes then also validate the direct measurement approach. This approach requires much less data than a revealed preference approach. So, once again, if one is able to fully specify a model for μ, or better still a model for μ and σ jointly, equivalence scales follow. Various attempts to specify such a complete model have been made (see, e.g., Kapteyn 1977; Kapteyn et al. 1980; Kapteyn et al. 1985; Kapteyn and Wansbeek 1985).

Since one can never be sure that a model is fully specified, the joint use of direct measurement and revealed preference allows for tests of specification that would not otherwise exist. In certain cases one can use the two different measurements to solve misspecification problems in a similar vein as in general latent variables models (see, e.g., Aigner et al. 1984).

Fourthly, it opens up new possibilities for identification. For example, if data series on consumption by households are too short to estimate all parameters in a demand system the availability of subjective measures, like μ, may help to identify parameters.

A Appendix

In order to obtain asymptotically efficient estimators of the parameters of interest (and hence a powerful test of the null), the length of the vector

$$g \equiv \frac{1}{N_c} Z_c' w + \frac{1}{N_s} Z_s' \mu \pi_4 - \frac{1}{N} Z' X_1 \theta - \frac{1}{N_c} Z_c' X_{2c} \pi_4 \ , \tag{A.1}$$

(cf. 7.4) has to be minimized in the appropriate metric, i.e. the inverse of the asymptotic variance covariance matrix of the vector g, where for all parameters true values have been inserted. An asymptotically equivalent procedure is to replace true parameter values by consistent estimates. These consistent estimates are obtained by minimizing g in the unit metric.

The derivation of the asymptotic variance covariance matrix of g is straightforward. For a start we assume that observations in two different samples are mutually independent. We can write

$$g \equiv g_c + g_s \equiv \left[\frac{1}{N_c} Z_c' w - \frac{1}{N} Z_c' X_{1c} \theta - \frac{1}{N_c} Z_c' X_{2c} \pi_4 \right] + \left[\frac{1}{N_s} Z_s' \mu \pi_4 - \frac{1}{N} Z_s' X_{1s} \theta \right] \ , \tag{A.2}$$

in obvious notation. Deote the asymptotic variance covariance matrices of g_c and g_s by Φ_c and Φ_s respectively. That is, Φ_i, $i = c, s$ is defined as the variance covariance matrix of the limiting distribution of $\sqrt{N_i} g_i$ for $N_i \to \infty$. Furthermore, let p_i be the limit for $N \to \infty$ of N_i/N, $i = c, s$. Let the asymptotic variance covariance matrix Φ of g be defined analogously to those of g_c and g_s, but with N_c or N_s replaced by N. Then the asymptotic variance covariance matrix of g is $(1/p_c)\Phi_c+(1/p_s)\Phi_s$. There is no need to derive Φ_c and Φ_s explicitly, one only needs to find consistent estimators that can be used in estimation. Define the vectors

$$g_{ci} = Z_{ci}\left(w_{ci}-\frac{N_c}{N}X'_{1ci}\theta-X'_{2ci}\pi_4\right)\equiv Z_{ci}eci \tag{A.3}$$

$$g_{si} = Z_{si}\left(\mu_{si}\pi_4-\frac{N_s}{N}X'_{1si}\theta\right)\equiv Z_{si}e_{si} \ , \tag{A.4}$$

where Z'_{ci} is the i-th row of $Z_c, X_{1ci}, X_{2ci}, Z_{si}, X_{1si}$ are defined analogously. The "residuals" e_{ci} and e_{si} are defined implicitly. Let the sample covariance matrices of g_{ci} and g_{si} be denoted as $\hat{\Phi}_c$ and $\hat{\Phi}_s$ respectively. These are consistent estimates of Φ_c and Φ_s. The estimate of Φ is

$$\hat{\Phi} = \frac{N}{N_c}\hat{\Phi}_c+\frac{N}{N_s}\hat{\Phi}_s \ . \tag{A.5}$$

In finite samples the variance covariance matrix of g is then approximated by $\Phi^* = \frac{1}{N}\hat{\Phi} = \frac{1}{N_c}\hat{\Phi}_c+\frac{1}{N_s}\hat{\Phi}_s$.

Using the notation of Sect. 7, cf. (7.6), let $W=[X_1,y]$, $\phi=(\theta',\pi_4)'$ then Φ^* is an estimate of the variance covariance of the error in the regression:

$$\frac{1}{N}Z'z = \frac{1}{N}Z'W\phi+\text{error} \ . \tag{A.6}$$

Efficient estimation amounts to GLS in this equation. Let $\sum_{ZW} = \text{plim}\,\frac{1}{N}Z'W$. Then the asymptotic variance covariance matrix of the estimator of ϕ is:

$$\text{avar}\,(\hat{\phi}) = (\textstyle\sum'_{ZW}\Phi^{-1}\sum_{ZW})^{-1} \ . \tag{A.7}$$

In finite samples the variance covariance of the estimator of ϕ is approximated by

$$\text{var}\,(\hat{\phi}) = \left[\frac{1}{N}W'Z\left(\frac{1}{N_c}\hat{\Phi}_c+\frac{1}{N_s}\hat{\Phi}_s\right)^{-1}\frac{1}{N}Z'W\right]^{-1} \ . \tag{A.8}$$

Thus, the computation of the efficient IV estimates amounts to the following procedure. First estimate Eq. (7.6) by IV assuming a scalar variance covariance matrix of the errors. Next form per observation the residual vector times the instrument vector (cf. (A.3) and (A.4)). Multiply these by N_c or N_s, depending on which subsample the observation belongs to. Compute the sample covariance matrices of these vectors per subsample, i.e. compute $\hat{\Phi}_c$ and $\hat{\Phi}_s$. Next form $\Phi^* = (1/N)\hat{\Phi}$, cf. (A.5). Use this result to perform GLS on (A.6), i.e. compute:

$$\hat{\phi} \equiv \left\{ \frac{1}{N} W' Z \hat{\Phi}^{-1} \frac{1}{N} Z' W \right\}^{-1} \frac{1}{N} W' Z \hat{\Phi}^{-1} \frac{1}{N} Z' z \ . \tag{A.9}$$

The variance covariance matrix of this estimator is then computed as

$$\text{var}\,(\hat{\phi}) = \left\{ \frac{1}{N} W' Z \hat{\Phi}^{-1} \frac{1}{N} Z' W \right\}^{-1} \ . \tag{A.10}$$

References

Aigner DJ, Hsiao C, Kapteyn A, Wansbeek TJ (1984) Latent variables in econometrics. Handbook of econometrics, vol II. North-Holland, Amsterdam, pp 1321–1393

Alessie RJM, Kapteyn A (1991) Habit formation, interdependent preferences and demographic effects in the almost ideal demand system. Econ J 101:404–419

Alessie RJM, Kapteyn A, Melenberg B (1992) Subjective definitions of poverty and the measurement of reference groups. Working paper Tilburg University

Angrist JD, Krueger AB (1992) The effect of age at school entry on educational attainment. An application of instrumental variables with moments from two samples. J Am Statist Assoc 87:328–336

Antonides G, Kapteyn A, Wansbeek TJ (1980) Reliability and validity assessment of ten methods for the measurement of individual welfare functions of income. Working paper Tilburg University

Arellano M, Meghir C (1988) Female labour supply and on-the-job search. An empirical model estimated using complementary data sets. Rev Econ Studies 59:537–559

Banks J, Blundell R, Lewbel A (1994) Quadratic engel curves, indirect tax reform and welfare measurement. Department of Economics, University College London, Discussion Paper 94-04

Bekker PA (1994) Alternative approximations to the distribution of instrumental variable estimators. Econometrica 62:657–685

Blackorby C, Donaldson D (1988) Adult-equivalence scales and the economic implementation of interpersonal comparisons of well-being. Discussion Paper 88-27, University of British Columbia

Blackorby C, Donaldson D (1993) Costs of children. Hand-out for presentation at the 1993 AEA meetings, Anaheim, California

Blundell R, Lewbel A (1991) The information content of equivalence scales. J Economet 50:49–68

Bound J, Jaeger DA, Baker RM (1993) The cure can be worse than the disease. A cautionary tale regarding instrumental variables. University of Michigan, working paper

Browning M (1992) Children and household economic behavior. J Econ Liter 30:1434–1475

Browning M (1993) More than one angel can dance on the head of a pin. Notes on adult equivalence scales. Paper presented at the 1993 AEA meetings, Anaheim, California

Buyze J (1982) The estimation of welfare levels of a cardinal utility function. Eur Econ Rev 17:325–332

Colasanto D, Kapteyn A, van der Gaag J (1984) Two subjective definitions of poverty. Results from the Wisconsin basic needs study. J Human Resources 28:127–138

Danziger S, van der Gaag J, Taussig M, Smolensky E (1984) The direct measurement of welfare levels. How much does it cost to make ends meet? Rev Econ Statist 66:500–505

Davidson R, MacKinnon JG (1993) Estimation and inference in econometrics. Oxford University Press, Oxford

Deaton AS, Muellbauer J (1980) An almost ideal demand system. Am Econ Rev 70:312–326

De Vos K, Garner T (1986) An evaluation of subjective poverty definitions. Comparing results from the US and the Netherlands. Rev Income Wealth 37:267–285

Dubnoff S (1979) Experiments in the use of survey data for the measurement of income minima. Working paper, Center for Survey Research, University of Massachusetts, Boston

Fisher FM (1987) Household equivalence scales and interpersonal comparisons. Rev Econ Studies 54:519–524

Hagenaars AJM (1986) The perception of poverty. North-Holland, Amsterdam

Hartog J (1988) Poverty and the measurement of individual welfare. J Human Resources 23:241–266

Kapteyn A (1977) A theory of preference formation. Unpublished PhD thesis, Leyden University. Leiden

Kapteyn A (1985) Utility and economics. De Economist 133:1 – 20

Kapteyn A, van Herwaarden FG (1980) Interdependent welfare functions and optimal income distribution. J Publ Econ 14:375 – 397

Kapteyn A, van Praag BMS (1976) A new approach to the construction of family equivalence scales. Eur Econ Rev 7:313 – 335

Kapteyn A, Wansbeek TJ (1982) Empirical evidence on preference formation. J Econ Psychol 2:137 – 149

Kapteyn A, Kooreman P, Willemse R (1988) Some methodological issues in the implementation of subjective poverty definitions. J Human Resources 23:222 – 242

Kapteyn A, van de Geer SA, van de Stadt H (1985) The impact of changes in income and family composition on subjective measures of well-being. In: David M, Smeeding T (eds) Horizontal equity, uncertainty, and economic well-being. University of Chicago Press, Chicago, pp 35 – 64

Kapteyn A, Wansbeek TJ, Buyze J (1980) The dynamics of preference formation. J Econ Behavior Organ 1:123 – 157

Lewbel A (1989) Household equivalence scales and welfare comparisons. J Publ Econ 39:377 – 391

Lusardi A (1993) Euler equations in micro data. Merging data from two samples. Center Discussion Paper 9304

Muellbauer J (1975) Aggregation, income distribution and consumer demand. Rev Econ Studies 42:525 – 543

Nelson JA (1993) Household equivalence scales. Theory versus policy? J Labor Econ 11:471 – 493

Pashardes P (1992) Equivalence scales in a rank-3 demand system. Working Paper IFS

Pollak RA, Wales TJ (1979) Welfare comparisons and equivalence scales. Am Econ Rev 69:216 – 221

Staiger D, Stock JH (1993) Asymptotics for instrumental variables regressions with weakly correlated instruments. Harvard University and NBER, working paper

Van Herwaarden FG, Kapteyn A (1981) Empirical comparison of the shape of welfare functions. Eur Econ Rev 15:261 – 286

Van Praag BMS (1968) Individual welfare functions and consumer behavior. North-Holland, Amsterdam

Van Praag BMS (1971) The welfare function of income in Belgium. An empirical investigation. Eur Econ Rev 2:337 – 369

Van Praag BMS (1991) Ordinal and cardinal utility. An integration of the two dimensions of the welfare concept. J Economet 50:69 – 89

Van Praag BMS, Kapteyn A (1973) Further evidence on the individual welfare function of income. An empirical investigation. Eur Econ Rev 4:32 – 62

Van Praag BMS, van der Sar NL (1988) Household cost functions and equivalence scales. J Human Resources 23:193 – 210

Van de Stadt H, Kapteyn A, van de Geer SA (1985) The relativity of utility. Evidence from panel data. Rev Econ Statist 67:179 – 187

Vaughan DR (1984) Using subjective assessments of income to estimate family equivalence scales. A report on work in progress. Proceedings of the Social Statistics Section, American Statistical Association 1984