

The Influence of Household Composition on Household Expenditure Patterns: Theory and Spanish Evidence

Angus S. Deaton

Princeton University

Javier Ruiz-Castillo

Instituto Nacional de Estadística, Madrid

Duncan Thomas

Yale University

A concept of demographic separability is proposed that formalizes the notion that there are groups of goods (adult goods) that have little or no relationship to specific classes of household demographics (the numbers or ages of children). That there exist adult goods demographically separable from children is a necessary but not sufficient condition for the validity of Rothbarth's method for measuring child costs. We propose two different methods for testing demographic separability and present results from a 1981 survey of Spain. The econometric evidence is in fair agreement with the theoretical presuppositions.

I. Introduction

This paper is concerned with two distinct but related areas of research. The first is the empirical study of the effects of household

This research was carried out while Javier Ruiz-Castillo was on leave at Princeton University on a Fulbright/MEC Fellowship. Duncan Thomas would like to thank the Sloan Foundation for financial support. We are grateful to Jim Heckman and Bob Pollak, whose suggestions had major effects on the shape of the paper, and to John Muellbauer, Ian Jewitt, and the referees of this *Journal* for numerous helpful comments.

composition on household consumption patterns and is one of the traditional concerns of household budget analysis. We estimate systems of Engel curves that incorporate the effects of household demographics using a large Spanish household survey. Our consumption data are comprehensive, and the survey distinguishes over six hundred categories of expenditure. It is therefore possible to select those goods for which we should expect there to be strong associations with particular demographic groups, as well as to distinguish groups of commodities for which such effects would be expected to be absent. In particular, we propose a concept of *demographic separability* that formalizes the idea that there are groups of goods with little or no relationship to a specific set of demographic variables. The most important example, and one for which we explicitly test, is the existence of adult goods that are demographically separable from children's characteristics.

Our second research area is the measurement of child costs: whether it can be done at all and, if so, what the relationship is between measurement of costs and the effects of children on the household budget. Our concept of demographic separability is an essential precondition for one important method of measuring the cost of children, that proposed originally by Rothbarth (1943). His method requires that there exist adult goods that are demographically separable from children. We show in this paper that this separability, while necessary for Rothbarth's procedure, is far from sufficient for its validity. Demographic separability is consistent with a number of different preference structures, and only one of these can be interpreted as supportive of the Rothbarth procedure. The basic problem is one that has been illuminated in recent discussions of child costs; see particularly Pollak and Wales (1979) and Deaton and Muellbauer (1986). Empirical evidence in and of itself is insufficient to identify the costs of children so that a variety of different theoretical structures, while observationally equivalent, can imply quite different estimates of costs. Demographic separability is an empirical proposition that is true or false on any given set of data, and so its truth cannot support any particular scheme for calculating child costs. However, its falsity would cast great doubt on the usefulness of Rothbarth's method. And since demographic separability is in itself an intuitively plausible concept, we believe that it is of considerable interest to see whether the evidence is for or against it.

The rest of the paper is presented as follows. Section II contains the basic theory. Demographic separability is introduced and defined, and we discuss the preference structures that are consistent with it. Section III proposes a practical method of testing for separability. We show how the effects of demographic variables on demand can conve-

niently be represented in terms of “outlay equivalent ratios,” quantities that measure the income equivalents of marginal changes in demographic characteristics. Demographic separability induces a simple proportionality between the outlay equivalent ratios for different goods, and we show how to test for this proportionality. Econometric issues are also discussed in this section. Section IV contains the empirical results; although it is easy to reject the separability restrictions overall, there is a good deal of conformity between the results and the predictions of the theory, and the concept of demographic separability appears to be useful for interpreting the data. Section V presents conclusions. We focus mainly on tests of separability, but for the benefit of those prepared to make the additional assumptions, we also use our results to calculate estimates of the costs of children.

II. Demographic Separability: Theory

To fix ideas, consider a traditional Engel curve model that is to be estimated on a single cross-sectional household survey for which it is assumed that all households face the same prices. Household expenditure on commodity i is written as

$$p_i q_i = g_i(x, \mathbf{a}, \mathbf{z}), \quad (1)$$

where $p_i q_i$ is expenditure on good i (note that prices and quantities are not separately observed), x is total expenditure (outlay or income) of the household, \mathbf{a} is a vector of demographic characteristics, and \mathbf{z} is a vector of other relevant covariates such as regions, seasonals, and occupational and educational dummies. The demographic vector \mathbf{a} can contain a wide range of information; in this paper we shall be concerned simply with the numbers of people in each of seven age categories. In consequence, in the empirical work, we shall take $a_j = n_j$, the number of people in the household who fall in the j th age group.

We wish to formalize the idea that some commodities are more closely connected than others with specific age groups in the household. Baby clothes are worn by babies, while (most) alcoholic drinks are consumed by adults. Consider the alcohol example further, and take one of its constituents, for example Bordeaux wine. What effect would we expect on the consumption of Bordeaux of the birth of an additional child into the household? Not much; Bordeaux is not like processed baby food or moviegoing in which the presence of an additional child would generate either a direct demand or else a rearrangement of the household budget to match the new circumstances. Even so, the effect is unlikely to be zero. Babies are not born with cash supplements, and there are more mouths to feed from the same total

budget, so that Bordeaux, like everything else, may have to take a cut. The effect of the child on Bordeaux consumption is essentially an income effect, while the effects on baby food or moviegoing also contain substitution effects that are absent for alcohol. In this case, we shall say that alcohol is *demographically separable* from children or from child demographic characteristics. It is easy to see that with only one such good, the separability restriction is not testable; *any* effect of children on alcohol consumption can be interpreted as an income effect. However, as soon as we have several potential adult goods, say Chateau Latour or Chateau Haut Brion, or Bordeaux and Rioja, then the effect of additional children on each ought to be proportional to the effects of changes in income on each. And that is a testable restriction.

We formalize these ideas quite generally, although in order to fix ideas and to lead into the empirical work, we use child demographics and adult goods as a running illustration. Nothing prevents the theory from being used to define demographic separability between, say, adult characteristics and child goods, with possible further application to the measurement not of child costs but of adult costs.

Start from a demographic group D —for example, children—but it could just as well be teenagers, grandparents, unrelated adults living in the household, and so on. Corresponding to D , we say that commodity group G , $G(D)$, is demographically separable from D if it is true that changes in the demographic structure within D exert only income effects on the goods in G . For G to be separable from D , we require that, for all g in G and all d in D ,

$$\frac{\partial q_g}{\partial a_d} = \theta_d \frac{\partial q_g}{\partial x}, \quad (2)$$

where the factor of proportionality θ_d is independent of the commodity g . Only the two derivatives in this expression are observable, so that testing the restriction requires a group with at least two goods in order to compute the ratios of the derivatives for different goods. Robert Pollak has pointed out to us that it is possible for there to be more than one demographically separable group of goods for any one demographic category. For example, for $d \in D$, there might be two groups G_1 and G_2 , with associated constants θ_{1d} and θ_{2d} in (2). In this case an additional child (say) causes a reallocation of resources between two groups of goods but changes patterns within the groups only insofar as the group expenditure totals are affected. We shall not consider this case further; our main concern is whether or not demographic effects can usefully be modeled as income effects, and the obvious starting point is that they do so for all unrelated, separable goods.

What happens for goods and demographic characteristics in which (2) does not hold? It is always possible to add a term to equation (2) and write, for all d and g ,

$$\frac{\partial q_g}{\partial a_d} = \theta_d \frac{\partial q_g}{\partial x} + \sigma_{gd}, \quad (3)$$

where θ_d has been calculated from some separable group satisfying (2) and σ_{gd} is defined by (3). This equation is chosen to look like the standard income and substitution decomposition of the effects of changes in price, and it is a useful analogy to think of demographically separable goods as those for which particular demographic changes have no substitution effects. However, for the analogy to be complete, it must be possible to interpret the quantity θ_d as the utility constant derivative of income with respect to the demographic characteristic a_d . As we shall see next, such an interpretation, while consistent with demographic separability, is not required by it, so that there exist preference structures for which the income and substitution interpretation would not be justified.

Consider a partition of the vector of demographics \mathbf{a} into $(\mathbf{a}_D, \mathbf{a}_D^*)$, where \mathbf{a}_D^* is that subvector of demographic characteristics *not* associated with D ; that is, it is the complement of \mathbf{a}_D in \mathbf{a} . Then the demographic separability condition (2) will hold if and only if the demand function takes the form

$$q_g = f_g[\phi(x, \mathbf{a}), \mathbf{a}_D^*] \quad (4)$$

for some function $\phi(\cdot, \cdot)$ and where the other covariates \mathbf{z} in (1) have been temporarily suppressed. While it is possible to provide a characterization of the preferences underlying (4), the results are not particularly enlightening. Instead, we consider two important classes of preferences that generate (4). Consider first a model of cost separability. Suppose that $c(u, \mathbf{p}, \mathbf{a})$ is the cost or expenditure function associated with utility level u , prices \mathbf{p} , and demographic characteristics \mathbf{a} and that the function takes the following form:

$$c(u, \mathbf{p}, \mathbf{a}) = c_1(u, \mathbf{p}_G, \mathbf{a}_D^*) + c_2(u, \mathbf{p}_C^*, \mathbf{a}_D). \quad (5)$$

The first part of the cost function, c_1 , contains the prices of the demographically separable goods G (adult goods) and the demographic characteristics that are *not* specifically associated with children, while the second part, c_2 , contains the demographic characteristics \mathbf{a}_D (child characteristics) and the prices of all goods *except* those of goods in G , that is, \mathbf{p}_C^* . Note how the various groups are constructed. We start from a demographic group D (children) and its associated characteristics \mathbf{a}_D ; from this a group of goods $G(D)$ (adult goods) is constructed. The complementary characteristics \mathbf{a}_D^* are those not in D (nonchild

characteristics) and the complementary goods G^* the non- G (non-adult) goods. Note the distinction between nonadult and “child” goods. There are many goods that are shared in the household (housing, heat, much food), and there are goods consumed by adults on which the presence of children may have major reallocative effects; these are nonadult goods but not in any sense child goods.

Cost-separable preferences imply compensated demand functions for the G -goods of the form, $g \in G$,

$$q_g = h_g(u, \mathbf{p}_G, \mathbf{a}_D^*) \quad (6)$$

so that adult goods depend only on child characteristics through the utility level u , and it is easily checked that both (2) and (4) hold. In particular, comparing (2) and (6), we have

$$\theta_d = \frac{\partial u / \partial a_d}{\partial u / \partial x} \quad (7)$$

so that θ_d has a natural interpretation as the change in outlay necessary to compensate for the change in a_d . It is this structure that can justify Rothbarth's (1943) method for calculating child costs. He proposed that costs be equated to the amount of money that would restore expenditure on adult goods to the level prevailing before the birth of the child. From (6), we can see that this suggestion is correct; at constant prices and nonchild characteristics, expenditure on adult goods indicates the level of utility u , at least if adult goods are normal, so that the function is monotonic.

Cost functions of the form (5) have previously appeared in the literature. Muellbauer (1976) proposes a model in which c_1 is the cost function for adults reflecting adult preferences and is defined over adult goods and adult characteristics, while c_2 is the corresponding cost function for child preferences. Adults are then assumed to allocate expenditure between themselves and their children so that utility is the same for both. Another simple specification of (5) is the form

$$c(u, \mathbf{p}, \mathbf{a}) = \sum_{k \in G^*} \gamma_k p_k + c_1(u, \mathbf{p}_G, \mathbf{a}_D^*), \quad (8)$$

which assumes that there are “required” quantities γ of child goods and that the adults buy these and then maximize their own utility with whatever resources remain. The cost function (8) is also consistent with Pollak and Wales's (1981) concept of “demographic translating” (see also Gorman 1976).

Cost separability is not the only preference structure that can be used to separate goods from demographics. An alternative and perhaps more obvious formulation is one that requires that the preference ordering over \mathbf{q}_G conditional on \mathbf{q}_G^* and \mathbf{a} be independent of both \mathbf{q}_G^* and the subvector \mathbf{a}_D^* . Under the usual representation conditions,

such independence requires that the utility function take the (weakly) separable form

$$u = u[v(\mathbf{q}_G, \mathbf{a}_D^*), \mathbf{q}_G^*, \mathbf{a}_D]. \quad (9)$$

As usual, weak separability implies the existence of subgroup demand functions for the separable group, so that for G -goods we now have

$$q_g = f_g(x_G, \mathbf{p}_G, \mathbf{a}_D^*), \quad (10)$$

where x_G is expenditure on G -goods in total. Since x_G is a function of prices, total outlay, and all demographic characteristics, (10) is consistent with both (4) and (2) so that weak separability, like cost separability, is consistent with demographic separability as originally defined. However, if we now use (10) to evaluate θ_d in (2), we get not (7) but

$$\theta_d = \frac{\partial x_G / \partial a_d}{\partial x_G / \partial x}, \quad (11)$$

and this is a useful measure of compensation only if we are prepared to assign direct welfare significance to total expenditure on G -goods. In general, there is little reason to do so. In the utility function (10), expenditure on adult goods is certainly a part of welfare, but it is not all of it. Adults may well have a separable subutility for a group of adult goods and activities, but that is not the same as being indifferent to other elements of (10), such as children's consumption levels or the amounts of goods that are shared.

Since the two separability concepts (5) and (9) both imply demographic separability and yet have such different welfare interpretations, it would be desirable to be able to separate between them. While this is possible in general, it is not possible on a single cross-sectional survey in which there is no price variation. Note first that from (6) we have

$$x_G = \sum p_g q_g = \sum p_g h_g(u, \mathbf{p}_G, \mathbf{a}_D^*), \quad (12)$$

so that if adult goods as a whole are normal, (12) can be inverted to give u as a function of x_G , \mathbf{p}_G , and \mathbf{a}_D^* . Substitution of this function back into (6) yields subgroup demand functions of the form (10), so that the existence of such demands is predicted by both approaches, a fact that we shall use in the empirical work below. To see the difference between the two approaches as well as the impossibility of detecting it with a single cross section, rewrite (6) as

$$q_g = h_g[\psi(x, \mathbf{p}, \mathbf{a}), \mathbf{p}_G, \mathbf{a}_D^*], \quad (13)$$

where $\psi(x, \mathbf{p}, \mathbf{a})$ is the indirect utility function. If $\chi_G(x, \mathbf{p}, \mathbf{a})$ is the function giving x_G , equations (10) become

$$q_g = f_g[\chi_G(x, \mathbf{p}, \mathbf{a}), \mathbf{p}_G, \mathbf{a}_D^*], \quad (14)$$

where the f_g functions are themselves utility consistent. With price variation, (13) and (14) are readily told apart; for example, the substitution effect between good g and a good outside G is zero in (13) and will generally be nonzero in (14). If the prices are suppressed from (13) and (14), they are empirically indistinguishable, at least if there is only one group G . If there are several distinct groups of goods that are each demographically separable from D , and we know in advance what these groups are, then under (14) the θ_d parameters will be indexed on G , while if (13) is true, θ_d has a unique welfare interpretation and should be the same for all groups.

In the empirical analysis below, we have a group of goods that are potential candidates for classification as adult goods, and we shall test whether they satisfy a single restriction like (2), that is, whether or not they are demographically separable from child characteristics. We have no prior basis on which to divide them into further groups, and any attempt to do so would only compromise our ability to construct even a single test. Our analysis therefore cannot shed light on whether or not it makes sense to calculate child costs should demographic separability be accepted on the data. Rothbarth's method of calculating child costs depends on whether or not expenditure on adult goods is thought to be a satisfactory indicator of adult welfare. Under (5) it is, under (9) it is not, and the choice is not something on which we can offer empirical evidence. But the very fact that demographic separability is a shared consequence of such diverse preference structures makes it more interesting to examine its relation to the data.

III. Model Formulation and Empirical Procedures

In our empirical analysis, we shall estimate Engel curves for a number of potential adult goods and then examine our results for evidence of demographic separability. The demographic characteristics that we consider are the numbers of people in each household that fall into seven distinct age ranges, from old people to babies. One convenient way to express the influence of demographic structure on consumption patterns is through what we call "outlay equivalent ratios." For any normal commodity i and any demographic category r , define the outlay equivalent ratio π_{ir} by

$$\pi_{ir} = \frac{\partial(p_i q_i)/\partial n_r}{\partial(p_i q_i)/\partial x} \cdot \frac{n}{x}. \quad (15)$$

Given any estimated Engel curve, these ratios can be calculated for each good and each demographic category, forming a matrix of

goods by categories. Each π_{ir} gives the effect of an additional person of type r on the demand for good i , measured as the amount of additional outlay that would have been necessary to produce the same effect on demand, that additional outlay expressed as a fraction of total household expenditure per household member. While the outlay equivalent ratios can be calculated for any combination of i and r , if good i belongs to a group G that is demographically separable from r , then, by (2), the π_{ir} coefficients will be the same for all goods in the group.

In order to calculate the outlay equivalent ratios, we require a functional form for the Engel curves. An ideal form for our purposes would be one for which demographic separability could be expressed in terms of parametric restrictions. We have devoted a good deal of time and energy to the search for such a form but have not been able to construct one. Instead, we select a flexible functional form for the Engel curve and accept the fact that we can examine demographic separability only at particular configurations of the explanatory variables. We use the following:

$$w_i = \frac{p_i q_i}{x} = \alpha_i + \beta_i \ln\left(\frac{x}{n}\right) + \eta_i \ln n + \sum_{j=1}^{J-1} \gamma_{ij} \left(\frac{n_j}{n}\right) + \delta_i \cdot \mathbf{z} + u_i. \quad (16)$$

The starting point for (16) is Working's (1943) Engel curve, which linearly relates the share of expenditure on each good to the logarithms of total expenditure. The effects of household composition are modeled by the inclusion of the logarithm of household size, $\ln n$, together with the ratios n_j/n to capture the additional effects of composition. In practice, outlay and household size tend to be more important than household composition, so that the demographic effects in (15) can be regarded as a linearization of the logarithm of a more general function of which n is the leading term. If $\eta_i = 0$, the demand for good i is unaffected by scaling household resources and household numbers, so that the sign patterns of η 's show how demand patterns change with household scale. The \mathbf{z} -variables, as before, are the other determinants of behavior.

We note finally that although (16) can be given a formal interpretation in utility theory, we choose not to emphasize it. Instead, we regard the equation as a convenient representation of the expectation of demand patterns conditional on the explanatory variables. For many of the goods we shall consider, substantial numbers of households record no purchases, so that the regression function is an average of zeros for households that do not purchase and of positive demands from those that do. As such, there is no straightforward interpretation of the regression function in terms of preferences. Our

tests of demographic separability should be interpreted with this in mind: we are testing whether the effects of demographics on expected purchases are analogous to the effects of income on expected purchases, where the effects include changes in demand at both intensive and extensive margins.

Given an Engel curve of the form (16), the outlay equivalent ratios can be calculated to be

$$\pi_{ir} = \frac{(\eta_i - \beta_i) + \gamma_{ir} - \sum_{j=1}^J \gamma_{ij}(n_j/n)}{\beta_i + w_i} \quad (17)$$

for $r = 1, \dots, J$, where γ_{ij} is defined to be zero. Estimates of the ratios are obtained by replacing the parameters by their estimates and replacing w_i and the n_j/n ratios by their values at the sample mean of the data. Having calculated the π 's, we test demographic separability for a group of v goods by testing, for $i = 1, 2, \dots, v$,

$$H_0: \Delta_{ir} = \pi_{ir} - \sum_j \frac{\pi_{jr}}{v} = 0. \quad (18)$$

We shall present estimates of the “discrepancies” Δ_{ir} , together with their asymptotic t -values, as well as a χ^2 -test for the hypothesis (18) as a whole. The details of the test statistics are given at the end of this section.

In order to provide a cross-check for our results, we exploit the theoretical structure of Section II to construct an alternative test of the separability hypothesis. For a group of normal goods G that is demographically separable from a demographic category D , equation (4) implies that there exist subgroup demand functions (cf. eq. [10] above)

$$q_g = f_g(x_G, \mathbf{a}_D^*) \quad (19)$$

and that these demand functions are independent of the demographic characteristics from the group D . Hence, for example, a test for demographic separability of a group of adult goods can be constructed by regressing each adult expenditure on total adult expenditure, on other variables \mathbf{z} , and on all demographic characteristics and by testing whether the child characteristics are jointly insignificant.

To match this alternative strategy to the first, we should have to use (16) to solve for total outlay x in terms of total expenditure on a specified group. The result could then be used to substitute out for x and to give a subgroup demand system containing all the demographic groups. However, the nonlinearities in (16) do not allow a closed-form solution, and the computations would be unnecessarily

complicated without one. Instead, we adopt a simple linear model, that is,

$$p_i q_i = b_{0i} + b_{1i} x_G + \sum c_{ij} n_j + \mathbf{d}_i \cdot \mathbf{z} + v_i. \quad (20)$$

The fact that (20) and (16) are mutually inconsistent is to some extent an advantage since the two sets of results will be different and there is the opportunity to check the robustness of the results to differences both in methodology and in functional form. Note finally that it is not appropriate to estimate (20) by ordinary least squares (OLS); x_G is defined to be the sum of the $p_i q_i$, so that it cannot be safely asserted that it is independent of any of the error terms v_i . The problem is easily dealt with by using instrumental variables with total expenditure x as the instrument for x_G .

The remainder of this section outlines the procedures for deriving standard errors and test statistics; the material can be skipped without loss of continuity. The starting point for the first set of experiments is to estimate equations (16) for a set of potentially separable goods. The appropriate technique is OLS; since the explanatory variables are the same in each equation, OLS is equivalent to system full-information maximum likelihood. Let $\hat{\mathbf{b}}_i$ denote the OLS estimate of the parameter vector from equation i , where \mathbf{b}_i is a vector containing the α , β , η , δ , and γ parameters from the i th equation. The variances and covariances of these parameter estimates are given by

$$E[(\hat{\mathbf{b}}_i - \mathbf{b}_i)(\hat{\mathbf{b}}_j - \mathbf{b}_j)'] = \sigma_{ij}(\mathbf{X}'\mathbf{X})^{-1}, \quad (21)$$

where \mathbf{X} is the matrix of common explanatory variables and σ_{ij} is the residual covariance between the i th and j th equations. If the vector of residuals from the i th equation is \mathbf{e}_i , then σ_{ij} is estimated from

$$\hat{\sigma}_{ij} = (n - k)^{-1} \mathbf{e}_i' \mathbf{e}_j, \quad (22)$$

where $n - k$ is the number of degrees of freedom in each regression.

Given the data, the π_{ir} 's are, from (17), nonlinear functions of the \mathbf{b}_i 's. Let \mathbf{J}_{ir} be the $1 \times k$ Jacobian matrix of the transformation from the \mathbf{b}_i 's into the scalar π_{ir} , so that the j th element of the matrix \mathbf{J}_{ir} is $\partial \pi_{ir} / \partial b_{ij}$. Then by the "delta method" (see, e.g., Fuller 1987, pp. 85–88), we have

$$\text{var}(\hat{\pi}_{ir}) = \sigma_{ii} \mathbf{J}_{ir}' (\mathbf{X}'\mathbf{X})^{-1} \mathbf{J}_{ir}. \quad (23)$$

With σ_{ii} replaced by its estimate from (22) and \mathbf{J}_{ir} evaluated at the parameter estimates and at the sample mean of the data, (23) is the formula used to generate the standard errors of the π 's reported in the next section. For the test of demographic separability corresponding to (18), define $\hat{\Delta}_{ir}$ by

$$\hat{\Delta}_{ir} = \hat{\pi}_{ir} - \sum_j \frac{\hat{\pi}_{jr}}{v}. \quad (24)$$

Let $\hat{\Delta}_r$ be the ν -vector of these discrepancies for the demographic category r , and $\hat{\pi}_r$ the corresponding π -ratios. If \mathbf{A} is the matrix $\mathbf{I} - (\mathbf{i}\mathbf{i}'/\nu)$ for identity matrix \mathbf{I} and vector of units \mathbf{i} , then (24) can be written as $\hat{\Delta}_r = \mathbf{A}\hat{\pi}_r$, while the generalization of (23) is

$$\{\mathbf{V}(\hat{\pi}_r)\}_{ij} = E[(\hat{\pi}_{ir} - \pi_{ir})(\hat{\pi}_{jr} - \pi_{jr})] = \mathbf{J}'_{ir}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{J}_{jr}\sigma_{ij}, \quad (25)$$

so that, under the null hypothesis, $\mathbf{V}(\hat{\Delta}_r)$ is $\mathbf{A}'\mathbf{V}(\hat{\pi}_r)\mathbf{A}$, which gives standard errors for the discrepancies. For the overall test for separability, note that if the true Δ_r are zero, the statistic

$$W_r = \hat{\pi}'_r \mathbf{A}' [\mathbf{A}' \mathbf{V}(\hat{\pi}_r) \mathbf{A}]^{-1} \mathbf{A} \hat{\pi}_r, \quad (26)$$

is asymptotically distributed as χ^2 with $\nu - 1$ degrees of freedom.

The procedures used in the second methodology are more familiar. Each equation is estimated by two-stage least squares (TSLS), and the absence of a subset of the variables is tested using an F -test, although, as with all inferences in TSLS, the distribution is only asymptotically valid. We also test the hypothesis that a particular demographic category, for example babies, enters none of the equations. Testing this cross-equation restriction requires the covariance between parameters in different equations, a covariance that is straightforwardly computed from the single-equation variance-covariance matrix and the estimated residual covariances from the TSLS estimates. The analogy is exact with equations (21) and (22) above for the OLS case.

IV. Data and Results

The data used in this study are taken from the Spanish *Encuestas de Presupuestos Familiares* and were collected during the year from April 1980 to March 1981. The main purpose of the survey, which has been repeated at irregular intervals over the last 20 years, is to provide the weights for the consumer price index. Since the survey is designed to calculate separate price indices for 50 provincial capitals, the sample size is large, 23,972 households from a total population of 10 million households. In this study we exclude the 264 households in the two North African cities of Ceuta and Melilla. Data were collected on 625 separate items of household expenditure. We shall use only a fraction of this detail, but its existence is necessary to allow us to pick out commodities that are specifically associated with individual age groups, even if the detail is then reaggregated. Expenditure data were collected over different reference periods for different commodities, so that food, drink, and tobacco (and many other items) were monitored over a 7-day period, while other expenditures were reported on a recall basis for periods of a month, 2 months, or a year, depending on the frequency with which the good is typically pur-

chased. We convert all expenditures to an annual basis, although two of the “lumpiest” items, purchases of motor vehicles and home repairs and improvement, were excluded from total household expenditure.

We also selected out a number of households on demographic grounds. Rather than expect the models (16) or (20) to apply to all household types, it seems wise to exclude those, such as single-person households, in which the life-style is sufficiently distinct to suggest that the effects of adding a child would be quite different from the marginal effect of an additional child in a household that already has children. We experimented with a number of different exclusion patterns. In the results reported below, all single-person households are excluded, 1,896 households, as well as those that report the presence of an unmarried couple without any children under the age of 14, a further 1,778 households. Our final, selected sample contains 19,951 households. We repeated the calculations with a more narrowly defined sample that also excluded 8,946 households in which there were no children under the age of 14. The results from this alternative data set differ only in detail from those reported below. We also “cleaned” data for individual households in which there were obvious inconsistencies. Although the cleaning was extremely time intensive, the total number of households involved was very small. Since the discovery of “unclean” data is inevitably (if unfortunately) a continuing process that proceeds along with the analysis, we obtained the sense that our estimates are quite robust to the cleaning, perhaps because we discovered no really egregious observations. Details of the cleaning operations will be made available on request.

We began by defining 12 possible adult goods, listed in table 1 and defined (where necessary) in the note to the table. A regression of the form (16) was estimated for the budget share of each good. The seven demographic categories are the numbers of people in the household aged 0–4 years, 5–8 years, 9–13 years, 14–17 years, 18–23 years, 24–60 years, and over 60 years of age. Also included in the regressions were 49 other variables representing the educational attainment of the household head, head’s and spouse’s ages and their squares, and dummies for multiple-earner households, head’s occupation, types of housing tenure (there is a good deal of rationing of subsidized housing in Spain), regional location, and date of interview.

The outlay equivalent ratios for the 12 goods are presented in the top panel of table 1 and their estimated standard errors are in the bottom panel. Since all the goods in the table are normal, a negative (positive) ratio implies that an additional person in the relevant age category acts like a decrease (increase) in total outlay. For a good and age group pairing in which there is demographic separability, addi-

TABLE 1
OUTLAY EQUIVALENT RATIOS FOR POSSIBLE ADULT GOODS

COMMODITY	AGE GROUP						
	60+	24-60	18-23	14-17	9-13	5-8	0-4
π -Coefficients							
Adult clothing	-.128	-.036	.131	.372	-.554	-.551	-.517
Adult education	1.045	1.043	2.596	3.299	.006	-.174	-.920
Alcohol	.027	.026	-.325	-.421	-.336	-.249	.292
Alcohol out	.630	1.004	1.377	.024	-.646	-.219	-.118
Entertainment	-.047	.205	.731	.796	-.388	-.387	-.602
Health	-.093	-.249	-.370	-.251	-.048	-.096	.520
Meals out	.020	.188	.129	-.013	-.437	-.232	-.230
Personal care	-.039	.056	.080	.059	-.518	-.236	-.136
Tobacco	.610	.975	1.650	.465	-.525	-.270	.060
Transport	-.026	.130	.049	-.293	-.274	-.096	-.158
Other (1)	-.269	-.153	.121	-.466	-.456	-.156	-.058
Other (2)	.055	.070	-.000	.059	-.209	-.131	-.363
All goods above	.012	.134	.243	.099	-.361	-.242	-.228
Standard Errors							
Adult clothing	.047	.041	.037	.044	.037	.039	.051
Adult education	.161	.145	.170	.197	.120	.124	.166
Alcohol	.094	.082	.073	.086	.074	.077	.102
Alcohol out	.113	.105	.102	.101	.087	.090	.118
Entertainment	.061	.054	.052	.059	.048	.050	.066
Health	.091	.079	.071	.084	.072	.075	.100
Meals out	.058	.052	.047	.054	.046	.048	.063
Personal care	.100	.088	.080	.093	.079	.082	.108
Tobacco	.114	.106	.108	.105	.088	.091	.120
Transport	.042	.037	.033	.038	.033	.034	.045
Other (1)	.089	.078	.072	.082	.071	.074	.097
Other (2)	.071	.062	.056	.066	.056	.058	.077
All goods above	.018	.015	.014	.016	.014	.014	.019

NOTE.—Adult clothing includes adult footwear. Alcohol out and meals out refer to alcohol and food consumed away from home. Other (1) and other (2) are collections of miscellaneous goods as follows: (1) insurance, gambling, funeral expenses, membership dues and subscriptions, licenses, taxes, and transfers to public institutions; (2) stationery excluding school materials, payments for financial services, and payments for personal services. The π -coefficients are calculated according to eq. (17) in the text.

tional people generate no direct demand for the good and the ratio must be negative. A group of goods all of which are demographically separable from a given age group will have the same negative π -ratios, at least up to sampling error.

The estimates in the table conform well to intuitive notions of the associations between goods and people. For all the goods taken together, additional adults (from 14 up) have positive effects on expenditures, and additional children (0-13) have negative effects (see the last row of the table). The left-hand side of the top panel of the table contains predominantly positive values. Adults have particularly

TABLE 2
OUTLAY EQUIVALENT RATIOS: DEVIATIONS FROM MEANS

COMMODITY	AGE GROUP		
	9-13	5-8	0-4
Adult clothing	-.188 (5.1)	-.318 (8.3)	-.331 (6.5)
Adult education	.371 (3.1)	.059 (.5)	-.734 (4.4)
Alcohol	.029 (.4)	-.016 (.2)	.478 (4.7)
Alcohol out	-.280 (3.2)	.014 (.2)	-.068 (.6)
Entertainment	-.023 (.5)	-.154 (3.1)	-.416 (6.3)
Health	.318 (4.4)	.137 (1.8)	.706 (7.1)
Meals out	-.072 (1.6)	.001 (.0)	-.044 (.7)
Personal care	-.153 (1.9)	-.003 (.0)	.050 (.5)
Tobacco	-.159 (1.8)	-.037 (.4)	.246 (2.0)
Transport	.092 (2.8)	.137 (4.0)	.028 (.6)
Other (1)	-.091 (1.3)	.077 (1.1)	.128 (1.3)
Other (2)	.156 (2.8)	.102 (1.8)	-.177 (2.3)
Wald test (11 df)	81.26	88.53	173.4

NOTE.—Absolute values of asymptotic *t*-statistics are in parentheses.

strong effects on adult education, on alcohol consumed away from home, and on tobacco. However, the most important feature of the table is the fact that, with only four exceptions (three for babies and one for 9-13-year-olds), all the outlay equivalent ratios for children are negative. The exceptions are themselves instructive. The positive effect of 9-13-year-olds on "adult" education is presumably an indication that the category has been too broadly defined; note the very large π -coefficient of 14-17-year-olds on the same category. It is not at all surprising that health expenditures are associated with the presence of babies. But there is also evidence that the presence of babies is associated with increased expenditure, presumably by the parents, on alcohol (consumed at home, not outside) and on tobacco. The alcohol effect is quite significant; the tobacco one not so. Such effects are consistent with Barten's (1964) model whereby the fact that adult goods do not have to be shared with children implies that their relative shadow prices fall as the number of children increases. Note too that the net effect on alcohol consumption, although positive, is smaller than that on alcohol consumed at home, a plausible consequence of babies making it more expensive to go out. Of course, the Barten model is not the only explanation of why infants cause their parents to drink and smoke more.

The results in table 1 suggest that demographic separability is worth more serious testing. Table 2 presents the relevant evidence on the equality of the π -coefficients. For each of the child demographic groups, the table shows the deviations of the π 's from their mean over the 12 goods, that is, the "discrepancies" Δ_{it} of equation (18), together

with the absolute values of their asymptotic t -statistics. The last row of the table shows the Wald tests for the hypothesis that all 12 goods are jointly demographically separable from the child category in that column; under the literal truth of the hypothesis, these test statistics should be distributed as chi-squared with 11 degrees of freedom. The high t -values occur where they might be expected given the results in table 1. For babies, health and alcohol π 's are significantly above the mean, while the π 's for adult clothing, education, and entertainment are significantly below it. For the 5–8-year-old children, there are fewer large t -values, although adult clothing and entertainment are significantly lower than the mean, and transport is significantly above it. For the older children, the π -values for adult education and health are clearly too large, and those for alcohol out and, again, adult clothing are too small. The overall Wald tests are highly significant at any conventional significance level, a result that is not at all surprising given the t -values in the columns above.

There are various interpretations that can be placed on these results, and rather than argue for one of them in particular, we prefer simply to catalog the possibilities. First, the standard errors, t -values, and χ^2 -statistics are all computed without allowance for residual heteroscedasticity in the underlying regressions. Although it is both expensive and complex to recalculate all the test statistics, we have followed White (1980) and Efron (1982) in estimating heteroscedastic-consistent variance-covariance matrices for the regression coefficients, for the π 's in table 1 and their discrepancies in table 2. The results are very much what would be expected from other similar studies. On average, the standard errors on the regression coefficients increase by about 20 percent and the t -values on the discrepancies by about 12 percent. We could therefore expect the separability test statistics to fall by about 25 percent, which is far from sufficient to bring the figures close to conventional acceptance levels.

Second, it is possible to question the relevance of conventional significance levels themselves. Our sample contains nearly 20,000 observations so that, if the size of the test is set at conventional levels, the probability of rejection is likely to be very much higher than it normally is, even if the hypothesis is, in some sense, very close to being true. In such circumstances, it may be better to try to balance explicitly the probabilities of type 1 and type 2 errors using a Bayesian approach. Schwarz (1978) has proposed a test criterion that asymptotically chooses the model with the higher posterior odds, and although his derivation makes an assumption of exponential families, the test applies more generally (see Chow 1983, pp. 300–302). In the current context, the Schwarz criterion is to accept the restricted model if the χ^2 -value is less than the logarithm of the sample size multiplied by the

number of restrictions, in this case a critical value of 108.9. By this criterion, we should accept that the 12 goods are demographically separable from the two oldest groups of children, but not from the youngest. To see whether this procedure also works when we expect to reject the null hypothesis, we calculated the Wald statistics for the adults, who should *not* be demographically separable from these goods. The values are 103.7 for those over 60, 188.7 for the 24–60-year-old group, 672.3 for the 18–23-year-old group, and 603.9 for the teenagers. Hence, apart from old people, who appear to get little in the way of these goods, we reject the (absurd) hypothesis that adults are separable from adult goods.

One can also view the results in table 2 as a “menu” for forming demographically separable subgroups. For example, if health is excluded from the babies group, as it clearly should be, the test value falls to 106.0, a figure that is (just) acceptable by the Schwarz criterion. Many other groupings are possible, and the choice is likely to be heavily influenced by prior conceptions of what is reasonable. For example, although adult clothing, adult education, and entertainment are three of the goods that have the largest discrepancies in the final column of table 2, a group containing only those three goods would all have similar π 's. It would be possible to tabulate Wald statistics for many such groupings, but quite apart from the risk of data mining, the results in the table give the essential information without further calculation.

Table 3 presents the TSLS estimates of important parameters from the subsystem demand equations. These results correspond to the model (20) above, which posits a linear relationship between expenditures on each good and total expenditure on adult goods, together with the demographic and other variables. The two-stage estimates are calculated using total expenditure as an instrument; OLS estimates are seriously biased by the simultaneity between the total adult expenditure and its components. Demographic separability implies that the characteristics of the separable demographic category should not appear in the regression, and this can be tested by the t -values for individual coefficients or by the F -statistics in table 4 (F_{c1} is the test for the exclusion of all three child groups, and F_{c2} is the test for the exclusion of the two older categories).

Tables 3 and 4 also show the coefficients and test statistics associated with adults. The F_a -statistics in table 4 correspond to F_{c1} and F_{c2} in the first two columns and test for the absence of all adult effects in the adult goods equations. Although we generally expect the presence of adults to affect expenditure on adult goods and demographic separability to be rejected, the estimates in table 3 should not be interpreted in the same way as the estimates of π -coefficients in table

TWO-STAGE LEAST SQUARES ESTIMATES OF ADULT GOOD SUBSYSTEM

Commodity	60+	24-60	18-23	14-17	9-13	5-8	0-4	mps*
Adult clothing	-2.80 (2.7)	-3.91 (4.4)	-.89 (1.2)	4.93 (6.2)	-2.71 (4.1)	-4.53 (6.3)	-3.59 (3.7)	.168 (79.6)
Adult education	5.13 (7.8)	3.13 (5.6)	9.20 (20.0)	10.68 (22.0)	1.13 (2.7)	.56 (1.3)	-2.72 (4.5)	.044 (33.7)
Alcohol	1.10 (3.0)	.97 (3.1)	-1.30 (5.0)	-1.16 (4.2)	-.07 (.3)	.25 (1.0)	1.65 (4.8)	.027 (36.6)
Alcohol out	1.32 (4.3)	3.57 (14.0)	4.05 (18.0)	.04 (.2)	-.41 (2.1)	.02 (.1)	.07 (.2)	.018 (29.5)
Entertainment	-2.57 (4.0)	.62 (1.2)	3.81 (8.4)	4.50 (9.4)	-.42 (1.0)	-1.04 (2.4)	-3.41 (5.8)	.078 (61.4)
Health	.80 (1.0)	-1.28 (2.0)	-2.75 (5.0)	-1.90 (3.3)	1.57 (3.3)	.95 (1.8)	4.36 (6.1)	.063 (41.1)
Meals out	-2.40 (2.6)	-.53 (.7)	-4.08 (6.1)	-2.72 (3.9)	-.92 (1.6)	-.53 (.8)	-.44 (.5)	.136 (72.3)
Personal care	-.01 (.0)	-.25 (1.2)	-.41 (2.4)	-.01 (.0)	-.39 (2.6)	.05 (.3)	.09 (.4)	.020 (40.8)
Tobacco	.65 (2.6)	1.77 (8.3)	3.05 (17.0)	.62 (3.3)	-.07 (.4)	.17 (1.0)	.86 (3.7)	.017 (35.1)
Transport	-.74 (.5)	-2.74 (2.2)	-9.45 (9.0)	-11.26 (10.0)	1.83 (2.0)	3.08 (3.1)	2.81 (2.1)	.312 (106.0)
Other (1)	-1.19 (1.6)	-1.64 (2.6)	-.86 (1.6)	-3.51 (6.1)	-.88 (1.9)	.27 (.5)	1.00 (1.4)	.068 (45.0)
Other (2)	.74 (1.4)	.29 (.7)	-.36 (1.0)	-.21 (.5)	1.35 (4.2)	.75 (2.1)	-.68 (1.4)	.048 (46.2)

NOTE.—Absolute *t*-values are in parentheses.

* mps is the marginal propensity to spend out of the total of the group so that the sum of the figures in the last column is unity.

TABLE 4
F-TESTS FOR DEMOGRAPHIC SEPARABILITY BY GOODS

Commodity	$F_{c1}(3)$	$F_{c2}(2)$	$F_a(4)$
Adult clothing	23.5	30.8	16.2
Adult education	10.9	4.9	223.7
Alcohol	8.0	.5	20.4
Alcohol out	1.6	2.3	113.7
Entertainment	12.8	3.7	58.2
Health	16.0	7.5	12.5
Meals out	1.2	1.7	14.2
Personal care	2.4	3.4	1.9
Tobacco	4.9	.6	86.3
Transport	5.9	7.3	51.5
Other (1)	2.0	1.8	11.0
Other (2)	9.1	11.8	1.2
1% critical value	3.8	4.6	3.3
0.1% critical value	5.4	6.9	4.6
Ln(sample size)	9.9	9.9	9.9

NOTE.— F_{c1} is the *F*-statistic for the exclusion of all three child variables from the regressions shown in table 3; F_{c2} is the test for the exclusion of the two older groups; F_a is the test for the exclusion of all adults.

TABLE 5
 χ^2 -TESTS FOR DEMOGRAPHIC SEPARABILITY
 BY DEMOGRAPHIC GROUP

	χ^2
Adults > 60	120.6
Adults 24-60	280.3
Adults 18-23	1,071.9
Adults 14-17	642.9
Children 9-13	67.8
Children 5-8	60.0
Children 0-4	136.2

1. Suppose that for some adult good all adult π -coefficients are positive and have the same value, so that additional adults, whatever their age, generate additional expenditures on the good in the same way as an increase in income would. Given this, the effects of those adults will be absorbed into the effect of total adult expenditure in the regressions in table 3, and the number of adults will have no effect either individually or jointly. Of course, there is no reason to expect this to happen since different-aged adults will typically have different effects on different adult goods. Even so, the insignificance of an adult category in the regression for an adult good does not mean that the good is not an adult good, only that adults of that age group have much the same effect on consumption as adults of other age groups.

The results in table 3 replicate the main features of table 1, although there are minor differences of detail; the choice of functional form and methodology does not appear to have a large effect on the results. All the child groups have significant negative coefficients in the adult clothing equation, so that, as before, children appear to depress expenditure on adult clothing more than they depress the expenditure on the other goods. There are again significant positive associations between the presence of babies and expenditures on health, alcohol, and tobacco and a significant negative coefficient in the entertainment and adult education equations, effects that would not be present were these four goods separable from babies. Once again, it is the baby group that generates the most conflict with separability; adult clothing apart, the older children have little effect on expenditures in the group once we control for total adult expenditure.

The test statistics in tables 4 and 5 provide a summary. In tests of whether each good is demographically separable from all three child groups, there are strong rejections for adult clothing, adult education, entertainment, and health. For all these goods, the F -test is larger than the Schwarz critical value, which, for an F -test, is the

logarithm of the sample size, or 9.9. At the conventional levels also shown, several other goods would fail. If babies are excluded, only the second “other” category fails the Schwarz test. If, instead of looking at goods, we test each child age group for separability with all 12 goods, we get the results shown in table 5. The results are similar to those for the original tests in table 2. At conventional significance levels, all the restrictions are rejected, but only babies fail the Schwarz criterion.

The results for the adults show the associations between particular age groups and particular goods. Tobacco consumption is associated with young adults in the 18–23 group, a group that also generates expenditures on education, entertainment, and alcohol consumed outside the home. Teenagers generate demand for clothing, education, and entertainment, while their negative coefficients for transport indicate not that they do not travel, but that their expenditure is less than that of other adults. The F -statistics in table 4 and the χ^2 -statistics in table 5 show the expected result, that adult composition effects are important for adult goods.

V. Conclusions

By conventional statistical criteria, nearly all the hypotheses with which we began have been strongly rejected with the Spanish data analyzed in this paper. Even so, we think that the concept of demographic separability with which we began is an interesting and useful one. Our selection of adult goods was a broad one, and several of the inclusions, such as expenditures on health, might have been expected not to be separable from children. Furthermore, we have a very large sample of nearly 20,000 households, so that conventional statistical criteria are not generous to even very small conflicts between theory and data. The large-sample Bayesian criterion used here presents a more plausible picture, that some of our provisional adult goods should not be regarded as such, but that most of the goods we tried are genuinely separable from children.

What then of the second of our topics, the measurement of child costs? Our theoretical analysis showed that even if demographic separability between adult goods and children were to be satisfied, the finding would not, of itself, justify the use of the Rothbarth method for measuring child costs. Of course, neither are we prohibited from going ahead, provided that we believe that expenditure on adult goods is a plausible indicator of adult welfare. On the supposition that it is worth following the consequences of the assumption, we conclude by presenting some estimates.

To do so, we return to the original functional form (16) and add

over a preselected group of adult goods, A , say, to get

$$w_A = \sum_{i \in A} \frac{p_i q_i}{x} = \alpha_A + \beta_A \ln\left(\frac{x}{n}\right) + \eta_A \ln n + \sum_{j=1}^J \gamma_{Aj} \left(\frac{n_j}{n}\right) + \delta_A \cdot \mathbf{z}, \quad (27)$$

where the subscript A denotes sums over the original indices i in (16). The right-hand side of (27), multiplied by x , yields an expression for total adult expenditure, which, according to Rothbarth, is monotonically related to adult welfare. We select a base household with two adults and no children, total expenditure x^0 , and sample mean characteristics and calculate its predicted expenditure on adult goods, x_A^0 . For some other household, for example, one with two adults and two babies, we calculate the value of total expenditure, x^1 , say, that would generate x_A^0 , that is, the same level of adult goods expenditure as in the base household. The cost of the additional children is then $x^1 - x^0$, which is, in general, a function of x^0 . Put differently, the two additional children are "equivalent" to $(x^1 - x^0)/x^0$ pairs of adults.

We use all 12 potentially adult goods to make this calculation. If the base household has the mean value of the logarithm of total household expenditure per capita, then an additional baby is calculated to be equivalent to 21 percent of an adult, an additional 5–8-year-old 22 percent of an adult, and an additional 9–13-year-old 31 percent of an adult. If education and health are dropped from the list, the costs are a little higher (as would be expected from the π -ratios) but increase to only 24 percent, 24 percent, and 35 percent, respectively. These figures are perhaps not implausible, but they are nevertheless subject to all the theoretical qualifications discussed above.

References

- Barten, Anton P. "Family Composition, Prices and Expenditure Patterns." In *Econometric Analysis for National Economic Planning*, edited by P. E. Hart, G. Mills, and J. K. Whitaker. London: Butterworths (for Colston Res. Soc.), 1964.
- Chow, Gregory C. *Econometrics*. New York: McGraw-Hill, 1983.
- Deaton, Angus S., and Muellbauer, John. "On Measuring Child Costs: With Applications to Poor Countries." *J.P.E.* 94 (August 1986): 720–44.
- Efron, Bradley. *The Jackknife, the Bootstrap, and Other Resampling Plans*. Philadelphia: Soc. Indus. and Appl. Math., 1982.
- Fuller, Wayne A. *Measurement Error Models*. New York: Wiley, 1987.
- Gorman, William M. "Tricks with Utility Functions." In *Essays in Economic Analysis*, edited by Michael J. Artis and A. R. Nobay. Cambridge: Cambridge Univ. Press, 1976.
- Muellbauer, John. "Can We Base Comparisons of Welfare between Households on Behavior?" Manuscript. London: Birkbeck Coll., Dept. Econ., 1976.

- Pollak, Robert A., and Wales, Terence J. "Welfare Comparisons and Equivalent Scales." *A.E.R. Papers and Proc.* 69 (May 1979): 216–21.
- . "Demographic Variables in Demand Analysis." *Econometrica* 49 (November 1981): 1533–51.
- Rothbarth, Erwin. "Note on a Method of Determining Equivalent Income for Families of Different Composition." App. 4 in *War-Time Pattern of Saving and Spending*, by Charles Madge. Occasional Paper no. 4. Cambridge: Cambridge Univ. Press (for Nat. Inst. Econ. and Soc. Res.), 1943.
- Schwarz, Gideon. "Estimating the Dimension of a Model." *Ann. Statist.* 6 (March 1978): 461–64.
- White, Halbert. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica* 48 (May 1980): 817–38.
- Working, Holbrook. "Statistical Laws of Family Expenditure." *J. American Statist. Assoc.* 38 (March 1943): 43–56.