Economics Letters 1 (1978) 93–98

THE DYNAMICS OF PREFERENCE FORMATION *

Arie KAPTEYN, Tom WANSBEEK and Jeannine BUYZE
Center for Research in Public Economics (CERPEC), University of Leyden, Leyden, The Netherlands

Received July 1978

Individual's preferences are explained on the basis of two types of influences, his own past consumption and the consumption of others which is directly observable by him. These effects are estimated using the "individual welfare function" approach of Van Praag, and a model of preferences formation.

1. Introduction

Individual preferences on consumption are usually supposed to be subject to two types of influences, viz. the individual's own past consumption (habit formation), and consumption by others (preference interdependence). In this paper [building upon earlier work, in particular Kapteyn, Van Praag, Van Herwaarden (1976), Kapteyn (1977)] we develop and estimate a model which quantifies both influences. We use a particular cardinal welfare function [the *individual welfare function*, developed by Van Praag (1968, 1971)] of which some distinctive traits will be sketched in the next section. After that the preference formation model will be presented. After discussing some econometric aspects the results are presented.

2. The individual welfare function

Following Lancaster (1966) and Van Praag (1968) we assume that individuals derive welfare from the characteristics of commodities rather than from commodi-

* The research reported in the paper is part of the Leyden Income Evaluation Project. It is supported by grants from the Netherlands Organization for the Advancement of Pure Research (ZWO) and the Dutch Ministry of Cultural Affairs, Recreation and Social Welfare. The empirical results are based on a survey, designed by B.M.S. van Praag, of members of the Dutch Consumer Union. Their willingness to make available the data for this study is gratefully acknowledged. Floor van Herwaarden kindly performed the tedious task of computing the reference weights used in the empirical analysis. The complete version of this paper is available under the same title as report 78.01 of CERPEC.

ties themselves. We define a *commodity group* as a set of commodities of which any combination of quantities can be described by the same (finite or infinite) set of characteristics.

Van Praag (1968) assumes that individuals are able to evaluate the welfare derived from a particular combination of characteristics by a number in the [0, 1]-interval. The number 0 will be given to the least preferred combination of characteristics and the number 1 to the most preferred combination. In general an individual will obtain combinations of characteristics by spending money on commodity groups. Assuming a one-to-one relation between the amount spent on a commodity group and the combination of characteristics thereby obtained, Van Praag infers that individuals will assign numbers from the [0, 1]-interval to money amounts spent on a commodity group. Under certain additional assumptions he finds that an individual evaluates an amount y spent on a commodity group according to a log-normal distribution function: $U(y) = \Lambda(y; \mu, \sigma)$. The parameters μ and σ may differ between individuals. The function $U(y)$ is called *the partial welfare function* (PWF) of the commodity group. In case the commodity group under consideration comprises all expenditure categories, i.e., y is total expenditures, or neglecting savings, after tax income, $U(y)$ is called the individual *welfare function of income* (WFI). A measurement procedure for U (i.e., of μ and σ per individual) has been developed by Van Praag (1971) and the lognormality of U has been extensively tested in a number of papers [1] with corroborative results. In the present study we use PWFs with respect to holiday expenditures of members of the Dutch Consumer Union who participated in a survey in 1971. The measurement results are presented in Kapteyn, Van Herwaarden, Van Praag (1977).

3. The formation of U

Since the PWFs of an individual represent his preferences, a preference formation theory has to explain the differences in the parameters μ and σ between individuals for various commodity groups. As our theory is the same for any commodity group we shall sketch the theory without continual references to a particular commodity group. When we speak of "expenditures" for example we implicitly mean "expenditures on the commodity group under consideration".

The basic idea of the preference formation theory is that an individual evaluates his expenditure by comparing it to the distribution of expenditures which he has perceived during his life-time. [2] If e.g. within his social reference group people mainly buy cheap refrigerators, he will evaluate a medium-priced one as excellent.

[1] Van Herwaarden, Kapteyn, Van Praag (1977) contains a short review.
[2] The theory is set out in detail by Kapteyn (1977).

Past habits have an important influence on the perceived expenditure distribution as well: if he previously bought expensive refrigerators, he will not easily be satisfied now by a medium-priced one. [3] Hence, the concept of a perceived expenditure distribution has to cover both the expenditures of others (preference interdependence) and the individual's own consumption habits up to the present (habit formation).

The concept of an expenditure distribution has been formalized by Kapteyn (1977). Let F_n denote the distribution function perceived by individual n and U_n his PWF. The preference formation theory now implies $U_n = F_n$.

As both U_n and F_n are distribution functions, their equality implies equality of their log-moments. Since μ_n and σ_n^2 are the first two log-moments of U_n, this yields relationships which have to explain the parameters μ_n and σ_n^2. In this paper we restrict ourselves to the explanation of μ_n. The definition of F_n given by Kapteyn (1977) implies the following explanation of μ_n:

$$\mu_n(0) = \beta_0 + \beta_1 \ln fs_n + (1-a) \sum_{t=-\infty}^{0} a^{-t} [\beta_2 \ln y_n(t) + \beta_3 \bar{m}_n(t)] + \epsilon_n(0) \; ; \qquad (1)$$

$\mu_n(0)$ is individual n's μ in year zero (the present); $\beta_0, \beta_1, a, \beta_2, \beta_3$ are parameters, with $\beta_2 + \beta_3 = 1$; $\epsilon_n(0)$ is a disturbance term, contemporaneously uncorrelated, with variance σ_ϵ^2; fs_n is individual n's family size; $y_n(t)$ is individual n's expenditure in year t; $\bar{m}_n(t)$ is the weighted average of log-expenditures as he perceives them in year t in his social reference group. The weights signify the relative importance of other individuals. For instance a close friend will get a large weight and a remote relative a small weight. The weights have been estimated by Kapteyn, Van Praag, Van Herwaarden (1976) and are inserted in the present analysis, to compute $\bar{m}_n$ per individual. The parameter β_2 represents the influence of preference interdependence, whereas the parameter β_3 represents the influence of habit formation.

4. Estimation

The data used in the estimation of (1) come from a survey of members of the Dutch Consumer Union, held in 1971. We only consider the commodity group "holiday expenditures" for which PWFs of 2081 individuals (i.e., their μ and σ) have been measured. Since the estimation of (1) requires in principle longitudinal data, which are unavailable, we employ an additional relation. Earlier research [Kapteyn, Wansbeek, Buyze (1977)] makes it plausible that individual n plans this year's expenditure $y_n(0)$ according to

$$\ln y_n(0) = \mu_n(-1) + \alpha\sigma_n(-1) + \zeta_n \; , \qquad (2)$$

[3] In the analysis prices are taken to be exogenous.

where α is a parameter and ζ_n an i.i.d. error term, with variance σ_ζ^2. Relation (2) makes it possible to remove almost all terms in (1) not pertaining to the present year. If we moreover assume $\sigma_n(0) = \sigma_n(-1)$, (1) and (2) imply:

$$\begin{aligned}\mu_n(0) = {} & (1-a)\beta_0 + (1-a)\beta_1 \ln fs_n + \{(1-a)\beta_2 + a\} \ln y_n(0) \\ & + (1-a)\beta_3 \bar{m}_n(0) - \alpha a \sigma_n(0) + \epsilon_n(0) - a(\epsilon_n(-1) + \zeta_n) \\ \equiv {} & \gamma_0 + \gamma_1 \ln fs_n + \gamma_2 \ln y_n(0) + \gamma_3 \bar{m}_n(0) + \gamma_4 \sigma_n(0) + u_n \,, \end{aligned} \tag{3}$$

where the γ's and u_n are implicitly defined. Eq. (3) can be estimated on the basis of the cross-section data available.

There are a few econometric problems left. First, the six parameters a, β_0, β_1, β_2, β_3 and α cannot be computed from the five γ's. Second, it is easily seen that $\ln y_n(0)$ correlates with u_n, which makes OLS-estimates of the γ's inconsistent. Third, as said before the $\bar{m}_n(0)$ are computed on the basis of estimates from earlier research. This induces measurement errors in $\bar{m}_n(0)$ (with assumed variance σ_δ^2) which also leads to inconsistency of the OLS-estimates of the γ's.

These three problems are solved by introducing a number of restrictions. First, with respect to α a prior distribution is specified based on earlier research [Kapteyn, Wansbeek, Buyze (1977)] with regard to 28 commodity groups (which did not include holiday expenditures). Second, $\beta_2 + \beta_3 = 1$ implies $\gamma_2 + \gamma_3 = 1$. Third, it is assumed that the population variances and covariances of μ and σ do not change over time (a kind of stationary assumption). This yields via (2) a restriction on σ_ζ^2.

The three restrictions are used to derive asymptotically unbiased parameter estimates. The estimation method employed is the so-called CALS-method [see Kapteyn, Wansbeek, Buyze (1978)].

5. Results

Table 1 presents the estimation results. The first two columns contain the estimates of the γ's by OLS and CALS, respectively. The third column contains the estimates of the β's and a.

Some conclusions:

(1) There is a considerable difference between the OLS estimates and the CALS estimates. Neglecting errors in variables, or more generally correlation between explanatory variables and the error term, apparently leads to estimates which may be wildly misleading.
(2) The value of σ_δ^2 (= 0.06) (with standard error 0.02) indicates that indeed the measurement errors of $\bar{m}_n$ could not be disregarded without introducing serious specification errors.
(3) The estimate of a is 0.55. This suggests that an individual "forgets" per year about half of his experiences with consumption patterns. The estimates of β_2

Table 1
Estimation results. [a]

Name of variable	OLS-estimates of regression coefficients	CALS-estimates of regression coefficients	CALS-estimates of original parameters
Constant	2.02	0.05	$\beta_0 = 0.10$
	(0.24)	(0.02)	(0.08)
$\ln fs_n$	0.15	0.12	$\beta_1 = 0.26$
	(0.02)	(0.02)	(0.11)
$\ln y_n$	0.42	0.69	$\beta_2 = 0.32$
	(0.01)	(0.03)	(0.14)
$\bar{m}_n$	0.31	0.31	$\beta_3 = 0.68$
	(0.04)	(0.03)	(0.14)
σ_n	−0.50	−0.45	$a = 0.55$
	(0.03)	(0.03)	(0.08)
Number of observations: 2081,		$\hat{\sigma}^2_\epsilon = 0.05$, (0.01)	$\sigma^2_\xi = 0.34$, (0.02) $\quad \sigma^2_\delta = 0.06$, (0.02)
OLS $\bar{R}^2$: [b]	0.49		
CALS $\bar{R}^2$: [c]	0.83		
	(0.04)		

[a] Asymptotic standard errors in parentheses. The computation method has been described by Kapteyn, Wansbeek, Buyze (1978).
[b] The corrected $\bar{R}^2$ obtained by performing OLS on (3).
[c] This quantity is defined as $1 - \hat{\sigma}^2_\epsilon/\text{var}(\mu_n)$, with $\text{var}(\mu_n)$ the sampling variance of μ.

and β_3 suggest that habit formation explains one third of an individual's preferences, whereas preference interdependence explains the remaining two thirds.

(4) The CALS $\bar{R}^2$ value suggests that only 17% in the variation of an individual's μ remains unexplained by model (1). For individual data this is a very low percentage. One should realize moreover that part of the unexplained variance of μ_n is due to measurement error in μ_n itself. Further improvement of the explanatory power of the model is possible, *inter alia* by choosing a more sophisticated specification of fs_n [cf. Kapteyn and Van Praag (1976)].

References

Kapteyn, A., 1977. A theory of preference formation, Unpublished Ph.D. thesis (Leyden University).

Kapteyn, A., F.G. Van Herwaarden and B.M.S. Van Praag, 1977, Direct measurement of welfare functions: Methods and results, Report 77.02 (Economic Institute of Leyden University, Leyden).

Kapteyn, A. and B.M.S. Van Praag, 1976, A new approach to the construction of family equivalence scales, European Economic Review 7, 313–335.

Kapteyn, A., B.M.S. Van Praag and F.G. Van Herwaarden, 1976, Individual welfare functions and social reference spaces, Report 76.01 (Economic Institute of Leyden University, Leyden).

Kapteyn, A., T.J. Wansbeek and J. Buyze, 1977, Maximizing or satisficing? Report 77.05 (Economic Institute of Leyden University, Leyden).

Kapteyn, A., T.J. Wansbeek and J. Buyze, 1978, Errors in variables: Consistent adjusted least squares (CALS) estimation, Report 78.04 (Center for Research in Public Economics, Leyden).

Lancaster, K.J., 1966, A new approach to consumer theory, Journal of Political Economy 74, 132–157.

Van Herwaarden, F.G., A. Kapteyn and B.M.S. Van Praag, 1977, Twelve thousand individual welfare functions, European Economic Review 9, 283–300.

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

Van Praag, B.M.S., 1971, The welfare function of income in Belgium: An empirical investigation, European Economic Review 2, 337–369.