98-05

Theodore MITRAKOS
Athens University of Economics and Business and Bank of Greece
and
Panos TSAKLOGLOU
Athens University of Economics and Business and IMOP

INCOME, CONSUMPTION OR SOMETHING ELSE?
SELECTING A WELFARE INDICATOR FOR DISTRIBUTIONAL STUDIES

Abstract

Empirical distributional studies usually rely on cross-sectional disposable income or consumption expenditure data. Due to short interview periods and extensive use of recall questions these variables exhibit artificially high variation, while due to life-cycle factors they do not exhibit a particularly high degree of correlation. Hence, they might not be considered as reliable indicators of the long-run welfare of the population members. The paper presents a simple methodology for extracting information about a more stable long-run welfare indicator from existing variables under assumptions that cannot be considered as particularly restrictive. An application is provided using the data of a Greek Household Budget Survey. The new indicator is more equally distributed than disposable income or consumption expenditure and is found to be more closely correlated with a number of non-monetary welfare indicators than the other two monetary welfare indicators.

Useful comments and suggestions on earlier versions of the paper were received from Andreas Andrikopoulos, Bruce Bradbury, Carlo Devillanova, Sofia Dimelis, Thesia Garner, Michael Magdalenos, Magda Mercader-Prats, Euclid Tsakalotos and conference participants at the Athens University of Economics and Business, the 51st Congress of the International Statistical Institute in Istanbul, the 1997 meeting of the Association of Southern European Economic Theorists (ASSET) in Marseilles and the 12th Annual Meeting of the European Society for Population Economics, in Amsterdam. The research of Theodore Mitrakos was support by a grant from the Research and Technology Ministry of Greece (grant no E341/96) and is part of his doctoral dissertation. The usual disclaimers apply.

Address for correspondence: Dr P. Tsakloglou, Department of International and European Economic Studies, Athens University of Economics and Business, 76 Patission Street, Athens 10434 GREECE. E-mail: panos@aueb.gr

1. Introduction

A common problem encountered in many empirical distributional studies is that of the selection of an appropriate distribution. Usually, when economists and other social scientists analyse inequality and poverty, they are ultimately interested in inequality in the distribution of welfare. However, welfare is not directly observable and hence, for the purposes of empirical studies, a reasonably close approximation to it has to be used instead. Standard microeconomic theory suggests that, other things being equal, an individual's welfare level is determined in the short-run by his/her levels of consumption and leisure and in the long-run by his/her level of "life-cycle" or "permanent" income. These notions of welfare are closely related to the concepts of "full income" and "earnings capacity" developed by Becker (1965) and Garfinkel and Haveman (1977). Reliable estimates of permanent income at the individual level can be obtained from long series of panel data. Very few such data sets exist in a small number of countries.

Regarding short-term concepts of welfare, since there are enormous difficulties in evaluating leisure in monetary terms, most empirical studies use current consumption or current income as welfare indicators. Each variable has its merits from a theoretical point of view. Current consumption is usually considered as a better approximation to life-cycle income than current income, because individuals and households tend to save and dissave in different periods of their life-cycles in an attempt to smooth their consumption and, thus, maximise their utility (assuming that utility is a positive but diminishing function of consumption). On the other hand, the use of current income has some advantages, since it can be considered as a better indicator of the ability of an individual or a household to achieve a particular welfare level [Sen (1992), Chaudhuri and Ravallion (1994)]. In practice, the data on consumption and income that are available for empirical studies in most countries usually come from Household Budget and Income Surveys and they are far from ideal.[1] Apart from being influenced by life-cycle factors, in most such surveys the relevant data are collected using extensively recall questions and are subject to large margins of error. As a result, in many instances the recorded level of correlation between income and consumption is

---

1. Note also that the data collected in such surveys concern consumption expenditures - not consumption. Although the two variables are closely related, they are not identical, the former being a “noisy” approximation of the latter (from a statistical point of view).

relatively low and a considerable proportion of the population who are classified as poor according to one welfare indicator appear close to the top of the distribution according to the other indicator. This finding may have disturbing implications for the design of policies aimed to alleviate poverty and/or reduce inequality, if the recorded levels of inequality and poverty as well as the composition of the poor or the structure of inequality are influenced by the welfare indicator used. Therefore, it is interesting to explore the possibility of constructing a composite indicator of the "permanent income" of the members of the population using the existing information about their current incomes and consumption expenditures. This is the aim of the present paper and examples are provided using the data of a Greek Household Budget Survey.

The remainder of the paper is organised as follows. The next section discusses briefly the data used, while the third section presents a methodology for the construction of an approximation of the "permanent income" of the members of the population. The fourth section is devoted to the analysis of inequality and poverty in Greece using alternative concepts of resources, while the final section concludes the paper and provides a discussion of its main findings.

2. The data

The paper uses the micro-data of the 1993/94 Greek Household Budget Survey which was carried out by the National Statistical Service of Greece. The survey covers all the non-institutional households of the country and its sampling fraction is 2/1000 (around 6,700 households or 20,000 individuals). It contains detailed information about consumption expenditures (actual and imputed), incomes after taxes, social security contributions and transfer payments, socio-economic characteristics of the households and their members as well as information on a number of housing amenities and consumer durables owned by the household. In order to approximate welfare as close as possible, the concepts of both current consumption expenditure and current income include, apart from actual consumption expenditures and net incomes, the value of consumption of income-in-kind evaluated at market prices. A number of adjustments were made to the data before they were used for the purposes of the paper. A few households were removed from the sample because the information they provided was considered to be extremely unreliable and the sample was re-weighted in order to reflect more accurately the entire population using weights

derived from the 1994 Labour Force Survey. Further, all consumption expenditure and income figures were expressed in constant mid-1994 prices in order to remove the impact of inflation (9.8% from the beginning to the end of the survey). Finally, the value of cars purchased during the period of the survey was subtracted from the concept of consumption expenditure and replaced by the value of imputed car services, estimated using hedonic regression techniques, for all the households which owned cars. Since around the time of the survey consumer credit was very limited in Greece – and, therefore, problems of double counting are likely to be very limited - the latter estimates were also added to the concept of income.

The unit of analysis is the household member and the corresponding distributions were normalised using the so called "modified OECD scales" [Hagenaars et al (1994)] which assign a weight of one to the household head, a weight of 0.5 to each of the remaining adults in the household and a weight of 0.3 to each child (person aged up to 13) in the household. Nevertheless, since the unit of information collection in the Household Budget Survey was the household, for the purposes of the derivation of the index of permanent income in the next section the unit of analysis is the household.[2]

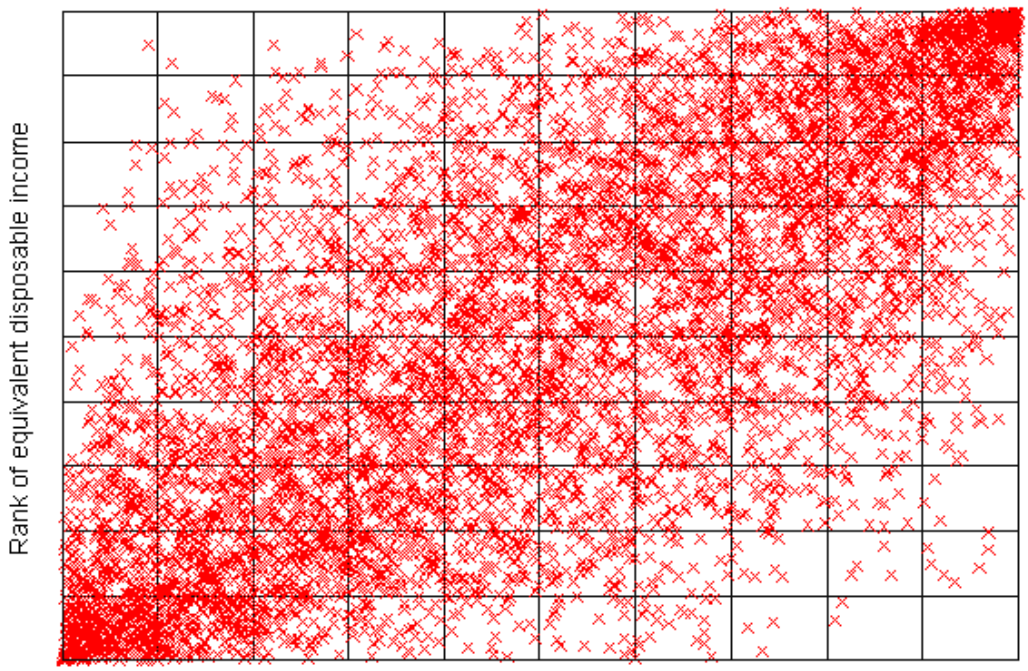
A casual first inspection of the distributions of equivalent consumption expenditure and equivalent income per capita reveals that they are relatively similar in terms of decile shares and inequality indices albeit, as anticipated, the former distribution is less unequal than the latter. For example, the share of the bottom (top) decile of the distribution of consumption expenditure is 3.4% (22.9%) while that of the bottom (top) decile of the income distribution is 3.1% (24.0%), the corresponding Gini indices being 0.295 and 0.310, respectively. However, a closer inspection of the data reveals that the two variables are not as closely related as one could anticipate. This is evident in Graph 1 and Table 1. In Graph 1 the members of the population are ranked from the least well-off to the most well-off, first, according to their equivalent consumption expenditure and, then, according to their equivalent income. The corresponding coordinates of these ranks are represented by one cross for each member, while the underlying grid splits the domain into deciles according to the relevant variable. Although most of the crosses are concentrated around the diagonal, a large number of them appear in the north west and south east areas of the graph. The Spearman rank

---

2. The empirical results of the next section are almost identical if the unit of analysis is the household member rather than the household (results available from the authors on request).

Graph 1. Scatterplot of ranks:

Equivalent consumption expenditure vs equivalent disposable income



correlation coefficient is 0.710. Table 1 summarises the evidence of Graph 1. In this table the members of the sample are ranked in deciles according to their equivalent consumption expenditure and equivalent income and then cross-tabulated.[3] Only 55.2% of them remain in the same or adjucent decile when moving from one distribution to the other, while almost a quarter of the sample (24.7%) move by three or more deciles.[4] There are even population members who belong to the top decile of one distribution and the bottom decile of the other. Hence, at least one of the two distributions cannot be considered as a good approximation of the unobservable distribution of "welfare". Part of these discrepancies should be attributed to genuine life-cycle factors, while another part should be attributed to the short interview period of the survey and the extensive use of recall questions. The latter is likely to add a lot

3. In Table 1 "0.0" denotes that the corresponding cell is non-empty but the proportion of the population in it is less than 0.05%.

4. The relatively low degree of correlation between the ranks of the members of the population in the distributions of consumption expenditure and income is not a peculiarly Greek phenomenon. See, for example, similar evidence for the U.K. and Spain cited in McGregor and Borooah (1992) and Mercader-Prats (1998), respectively. Similar evidence but in a slightly

Table 1. Cross-tabulation of population members ranked according to equivalent income and equivalent consumption expenditure per capita (% of the total population)

		Income Decile									
		1	2	3	4	5	6	7	8	9	10
Consumption expenditure Decile	1	4.9	2.2	1.3	0.4	0.5	0.3	0.2	0.1	0.1	
	2	2.0	2.4	1.8	1.4	1.0	0.6	0.5	0.3	0.0	0.0
	3	1.3	1.8	1.9	1.3	1.4	1.0	0.6	0.6	0.2	0.0
	4	0.7	1.4	1.5	1.6	1.2	1.2	1.0	0.7	0.3	0.3
	5	0.5	0.8	1.3	1.6	1.5	1.4	1.1	0.9	0.7	0.3
	6	0.2	0.6	0.8	1.5	1.5	1.4	1.5	1.0	1.0	0.3
	7	0.2	0.4	0.6	0.9	1.0	1.5	1.7	1.7	1.5	0.6
	8	0.1	0.3	0.4	0.7	0.8	1.4	1.5	1.8	2.1	1.0
	9	0.0	0.2	0.3	0.4	0.6	0.9	1.1	1.7	2.2	2.6
	10	0.0		0.1	0.1	0.3	0.4	0.9	1.4	2.0	4.9

of "artificial" variation to the estimates of both consumption expenditure and income. Under these circumstances it is worth-trying to construct a less "noisy" welfare indicator.

3. Permanent income

Following Abul Naga (1994), Abul Naga and Burgess (1997) and Mercader-Prats (1998), let X be the vector of available welfare indicators $[x_1, x_2, ..., x_k]'$, such as current income, consumption expenditure, etc. Further, assume that these indicators are related to the permanent income, y_p, (the "true" welfare indicator), in the following way:

---

different framework can be found in Anand and Harris (1994) for Sri Lanka and Hagenaars et al

$$X = by_p + U \tag{1}$$

where $b = [b_1, b_2, ..., b_k]'$ and $U = [u_1, u_2, ..., u_k]'$ is the vector of transitory components (error terms). (1) is a factor analysis model where y_p is not observable. A number of techniques can be used for the estimation of such models (method of moments, factor analysis, principal component analysis, etc). The choice of estimation technique depends on the number of welfare indicators available (k), as well as the number of additional assumptions that the researcher is willing to make.[5] Once the structural parameters of the system have been estimated and in order to extract information about y_p, additional assumptions have to be made about the joint distribution of X and y_p.

In this paper we assume that $y_p \sim N(\mu_p, \sigma_p^2)$ and $U \sim N(0, \Omega)$. In this case, from the properties of the normal distribution[6] (1) implies that $X \sim N(b\mu_p, bb'\sigma_p^2 + \Omega)$. Following Greene (1993, p. 76), the conditional distribution of permanent income y_p given the vector X, $f(y_p|X)$, will be:

$$y_p|X \sim N[\mu_p + \Sigma_{y_pX}\Sigma_{XX}^{-1}(X - b\mu_p), \quad \Sigma_{y_py_p} - \Sigma_{y_pX}\Sigma_{XX}^{-1}\Sigma_{Xy_p}] \tag{2}$$

where Σ is the $[(k+1)\times(k+1)]$ covariance matrix of $y_p, x_1, x_2, ..., x_k$, which can be broken down into the sub-matrices: $\Sigma_{y_py_p} = \text{cov}(y_p, y_p) = \sigma_p^2$, Σ_{y_pX}, Σ_{Xy_p} which are the $(1\times k)$ and $(k\times 1)$ covariance matrices of permanent income with $x_1, x_2, ..., x_k$, and Σ_{XX} which is the $(k\times k)$ matrix of covariances of $x_1, x_2, ..., x_k$. Then, permanent income is defined as:

$$E(y_p|X) = \mu_p + \Sigma_{y_pX}\Sigma_{XX}^{-1}(X - b\mu_p) \tag{3}$$

(1994) for twelve member-states of the European Union.

5. As Abul Naga (1994) points out, if $k > 3$, the model is over-identified and its estimation using the method of moments requires the imposition of additional assumptions (structure). Nevertheless, the advantage of the method of moments is that the estimated parameters are consistent and independent of the type of distribution of the welfare indicators X.

6. If $X \sim N(\mu, \sigma^2)$, then $a + bX \sim N(a + b\mu, b^2\sigma^2)$.

that is, the permanent income of each household is a linear function of all the available welfare indicators X of the household in question.[7] The weights assigned to the various welfare indicators are determined by the degree of covariance of these indicators both with permanent income and between themselves.

In order to derive an index of permanent income from our data, we assume that for every household in the sample the logarithms of their (current) income, Y, and consumption expenditure on non-durable goods,[8] C, are related with the logarithm of their permanent income, Y_p, in the following way:

$$Y = Y_p + Y_t \tag{4}$$

$$C = B + Y_p + C_t \tag{5}$$

where Y_t and C_t are, respectively, the transitory component of income and the error term of the consumption function (or, alternatively, transitory component of consumption). Both variables are used in logarithmic form since, using appropriate tests (x^2, Kolmogorov-Smirnov) it was found that consumption expenditure is approximately lognormally distributed, whereas in the case of current income the assumption of lognormality was only marginally rejected at the 5% level of significance. For the transitory components it is assumed that they have zero means and, further, that they are uncorrelated both with each other and with the corresponding permanent components:

$$\text{cov}(Y_p, Y_t) = \text{cov}(C_p, C_t) = \text{cov}(Y_t, C_t) = 0 \tag{6}$$

The first two assumptions are pretty innocuous, but this is not necessarily the case regarding the third assumption, $\text{cov}(Y_t, C_t)=0$, although this assumption is frequently made in macroeconomic studies. It implies that unanticipated changes in the current income of a household affect its current consumption only through their effect on

---

7. Bartholomew (1984) demonstrates that such an index can be constructed if the distribution of at least $k-1$ of the X indicators belongs to the family of exponential distributions (normal, gamma, Poisson, etc.).

8. The assumption that permanent income determines consumption expenditure on non-durable goods and services rather than total consumption expenditure is both theoretically appealing and commonly used in macroeconomics; see, for example, Hall (1978), Falk and Lee (1990) and Patterson (1992). Nevertheless, it should be noted that using total consumption expenditure instead of consumption expenditure on non-durable goods affects the results only marginally.

permanent income. In an attempt to relax this assumption, we estimated the covariance of the transitory components of disposable income and consumption expenditure on non-durable goods using the corresponding macroeconomic time-series of Dimelis et al (1997) after normalising the means of these distributions to the means of the relevant distributions of the 1993/94 Household Budget Survey.[9] However, the estimated covariance of Y_t and C_t was found to be very close to zero. For this reason, it was decided to stick to the assumption that $\text{cov}(Y_t, C_t)=0$.[10]

Taking (6) into account, the sample moments of (4) and (5) are:

$$\text{var}(Y)=\sigma_P^2+u_Y \tag{7}$$

$$\text{var}(C)=\sigma_P^2+u_C \tag{8}$$

$$\text{cov}(Y, C)=\sigma_P^2 \tag{9}$$

The system of these three equations can be identified and, hence, we can estimate the three unknown variances of permanent income σ_P^2, transitory income, u_Y, and transitory consumption, u_C. Estimates of the corresponding parameters are provided in Table 2. As anticipated, the proportional contribution of transitory income to the variance of disposable income (32.2%) is higher than the proportional contribution of the transitory consumption to the variance of consumption expenditure on non-durable goods (22.1%).

In our case the general model (1) as specified in equations (4) and (5), gives the following expression for (3):

$$E(y_P|X) = \mu_Y + \Sigma_{y_P X}\Sigma_{XX}^{-1}(X-\mu_X)$$

$$= \mu_Y + \frac{\sigma_P^2}{(\sigma_P^2+u_Y)(\sigma_P^2+u_C)-(\sigma_P^2)^2}[u_C(Y-\mu_Y)+u_Y(C-\mu_C)] \tag{10},$$

9. Dimelis et al (1997) derive transitory components of disposable income and consumption expenditure on non-durable goods by applying the Hodrick and Prescott (1997) filter on National Accounts data for the period 1960-94.

10. More specifically, the estimated value of the covariance of the two transitory components was equal to 0.000809. Estimates of permanent income derived using this non-zero covariance estimate were very similar to those reported below (results available from the authors on request).

Table 2. Transitory and permanent components of the variances of the logarithms of disposable income and consumption expenditure on non-durable goods

	Variance		
	Total	Permanent component	Transitory component
Disposable income	0.343	0.232	0.110
Contribution (%)	100.0	67.8	32.2
Consumption expenditure	0.298	0.232	0.066
Contribution (%)	100.0	77.9	22.1

that is, the permanent income of a particular household is equal to the mean of the disposable income of the entire population plus the weighted sum of the deviations of disposable income and consumption expenditure on non-durables of the household from the corresponding sample means. The weights depend positively on the variance of the transitory component of the opposite variable; in other words, the "noisier" one variable is the higher the weight assigned to the other variable. Finally, substituting the estimated values of the parameters σ_P^2, u_Y, and u_C in equation (10) we obtain the following:

$$E(Y_P|Y,C) = 1.850 + 0.317\,Y + 0.532\,C \qquad (11).$$

As anticipated, permanent income is found to be more closely related with the less "noisy" consumption expenditure and, therefore, its estimate is determined to a larger extent by this variable than by disposable income.

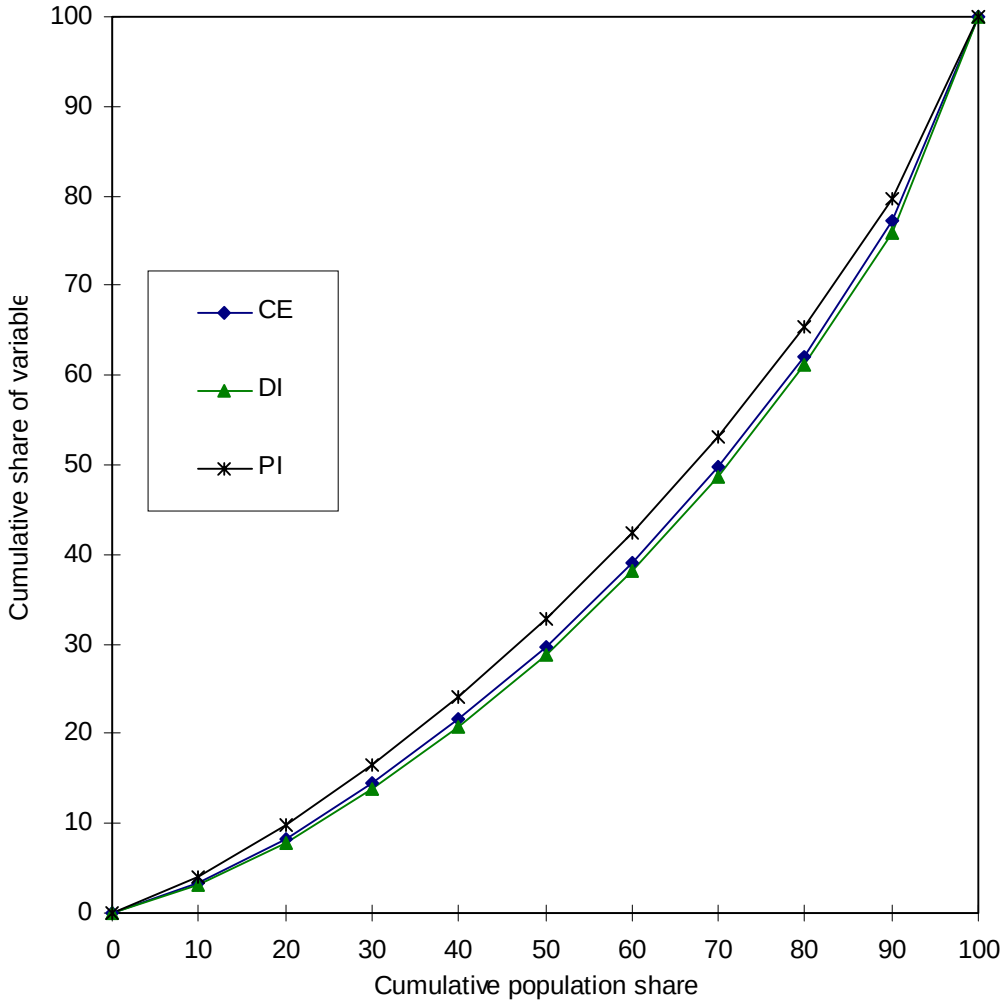
How does the new distribution compare with the distributions of disposable income and consumption expenditure? An answer to this question is provided in Table 3 and Graph 2. Table 3 provides decile shares and estimates of six widely used inequality indices for the distributions of equivalent income, equivalent consumption expenditure and equivalent permanent income per capita. Graph 2 reports the corresponding Lorenz curves. As one would anticipate, the new distribution appears to be far more equal than the other distributions and the Lorenz curve of the

Table 3. Decile shares and aggregate inequality indices

Population Decile / Index of inequality	Disposable Income (DI)	Consumption Expenditure (CE)	Permanent Income (PI)
1 (bottom)	3.1	3.4	4.3
2	4.8	5.0	5.9
3	5.9	6.1	6.9
4	7.0	7.2	7.9
5	8.1	8.2	8.7
6	9.3	9.3	9.7
7	10.6	10.6	10.8
8	12.3	12.4	12.1
9	14.9	15.0	14.1
10 (top)	24.0	22.9	19.5
Gini	0.310	0.295	0.234
Atkinson (ε =0.5)	0.080	0.070	0.044
Atkinson (ε =2.0)	0.280	0.255	0.166
Theil	0.170	0.145	0.090
Mean log deviation	0.163	0.145	0.090
Variance of logs	0.322	0.292	0.181

distribution of permanent income clearly dominates the Lorenz curves of the other distributions. Depending on the index and the distribution, the estimates of the inequality indices decline between 20% and 45% when moving to the last column of Table 3. Since the most important differences between the distribution of permanent income and the other distributions concern the shares of the top and bottom deciles, the largest proportional declines in inequality are recorded by those indices which are relatively more sensitive to changes in the tails of the distribution rather than the Gini

Graph 2. Lorenz curves for three distributions



CE: Equivalent consumption expenditure
DI: Equivalent disposable income
PI: Equivalent permanent income

index which is relatively more sensitive to changes around the median (Cowell (1995)). Further, these indices suggest that inequality is lower in the distribution of consumption than in the distribution of disposable income. In fact, the Lorenz curves reported in Graph 2 do not intersect and, thus, provide a complete rank of the distributions under examination.

The estimates of Table 3 were derived under the assumption that $\text{cov}(Y_t, C_t) = 0$. It may be argued that the value of $\text{cov}(Y_t, C_t)$ is likely to be positive and low but, nevertheless, higher than the estimate derived using macro-economic data (0.000809). As Mercader-Prats (1998) points out, if $\text{cov}(Y_t, C_t) = u_{YC} \neq 0$, then (10) should be modified to

$$E(Y_P|X) = \mu_Y + \Sigma_{Y_P X}\Sigma_{XX}^{-1}(X - \mu_X)$$

$$= \mu_Y + \frac{\sigma_P^2}{\sigma_P^2(u_Y + u_C - 2u_{YC}) + u_Y u_C - u_{YC}^2}[u_C - u_{YC})(Y - \mu_Y) + (u_Y - u_{YC})(C - \mu_C)] \quad (12)$$

Experimentation with alternative values of $\text{cov}(Y_t, C_t)$ showed that the estimates of Table 3 are rather robust. For example, the estimates of the Gini index of inequality for the distribution of permanent income under alternative assumptions regarding the value of $\text{cov}(Y_t, C_t)$ are the following: G_0=0.234 (as in Table 3), $G_{0.000809}$=0.233 (resulting from the macroeconomic estimate of $\text{cov}(Y_t, C_t)$) and $G_{0.005}$=0.229. Even when we assumed that $\text{cov}(Y_t, C_t)$=0.01 - that is, over twelve times higher than the corresponding macroeconomic estimate – the value of the Gine index changed by less than 5%; $G_{0.01}$=0.224.

Is the new variable able to predict the relative welfare position of the population members better than the existing variables? An attempt to provide an answer to these questions is provided in Tables 4 and 5. As noted in section 2, the Household Budget Survey contains information on a number of housing amenities and consumer durable goods of each household. In Table 4, three such items with wide variation across welfare levels are selected: second (holiday) home, car and dishwasher. Then, the population is grouped into deciles according to each of the three welfare indicators and the proportion of each decile living in households with such items is estimated. Naturally, the proportion of each decile with access to such items increases as we move from the bottom to the top of the distribution, irrespective of the welfare indicator used. Although the differences across indicators are not very large, permanent income appears to perform better than the other indicators. The bottom line of Table 4 reports the coefficient of variation of the above decile shares as an, admittedly very crude, indicator of the "goodness of fit". Roughly speaking, the higher the value of the coefficient of variation for each particular item, the better the corresponding indicator approaches the welfare ranking of the population according

Table 4. Members of deciles ranked according to alternative welfare indicators with access to particular housing amenities and consumer durables

Decile	Second home			Car			Dishwasher		
	DI	CE	PI	DI	CE	PI	DI	CE	PI
1	1.2	0.7	0.6	12.8	9.1	6.7	3.9	1.6	2.7
2	3.9	4.2	2.8	22.7	21.2	21.5	5.8	5.7	4.5
3	6.0	3.9	4.9	39.9	35.1	31.9	7.6	9.0	7.2
4	6.6	7.8	6.3	41.8	47.5	44.8	11.0	11.5	8.1
5	7.9	9.3	10.3	48.9	53.7	48.7	14.9	14.8	12.8
6	10.3	11.6	10.6	57.8	55.3	61.1	15.2	18.5	19.8
7	13.9	14.0	12.1	63.9	66.9	65.	22.4	22.2	23.5
8	13.9	14.8	14.8	70.6	71.2	72.1	30.7	27.0	27.6
9	20.4	17.9	20.4	78.2	77.7	82.8	35.6	36.6	40.1
10	24.3	24.2	25.4	85.7	84.5	86.9	51.2	51.6	52.2
CV	0.678	0.661	0.720	0.451	0.470	0.506	0.766	0.769	0.825

to the corresponding item.[11] In all cases permanent income appears to correlate better with the welfare status of the population members as it is reflected in the three items used there although, once again, the differences across indicators are not very large.

The situation is more clear in Table 5. For the purposes of this table three new indices are constructed. The first index (INDEX1) exploits the information on housing amenities available in the Household Budget Survey. For each population member the value of the index is the average score on seven items, the weights of the items being the proportion of the population living in households with such amenities. These amenities and the corresponding proportion of the population living in households with such items (in parentheses) are: dwelling with bath or shower (93.4%), dwelling with hot water (92.4%), separate kitchen inside the dwelling (97.5%), dwelling with telephone (89.2%), WC inside the dwelling (91.7%), at least 40 square meters available per equivalent adult in the household (68.2%) and second (holiday) home (10.8%). The

11. For example, if according to a particular indicator the proportion of all deciles living in households with cars was the same, the value of the coefficient of variation would be equal to zero. Similar results were obtained when other housing amenities and consumer durables with lower variation across population deciles were used instead of second home, car and dishwasher.

Table 5. Spearman rank correlation coefficients of alternative welfare indicators

	DI	CE	PI	INDEX1	INDEX2	INDEX3
DI	1.000					
CE	0.695	1.000				
PI	0.879	0.942	1.000			
INDEX1	0.426	0.409	0.450	1.000		
INDEX2	0.509	0.523	0.562	0.416	1.000	
INDEX3	0.523	0.536	0.579	0.399	0.945	1.000

second index (INDEX2) is the counterpart of INDEX1 for consumer durable goods. The following nine items were selected: refrigerator (98.6%), electric cooker (79.0%), vacuum cleaner (65.5%), colour TV (88.1%), video (41.9%), hi-fi (37.4%), washing machine (85.8%), dishwasher (19.8%) and car (52.2%). Since the cost for obtaining these items varies considerably across items and information on the average cost per item exists in the Household Budget Survey, it was decided to construct a third index (INDEX3) reflecting the average monetary value of the corresponding stock of durable goods for each household.[12]

Once the scores for every member according to each of the three indices were calculated, the population was ordered from the member with the lowest to the member with the highest score according to each index and the corresponding ranks were estimated. The Spearman rank correlation coefficients of each of these indices and the three monetary indicators are reported in Table 5. In all cases the correlation of the ranks of the population members according to their permanent income with the ranks according to any of the three indices are substantially higher than the corresponding correlation coefficients of the ranks of the other monetary indicators (disposable income and consumption expenditure) with the ranks of these indices.

12. It should be stressed that, for a number of reasons, these indices should not be considered as better indicators of the standard of living than the monetary indicators and they are used here for illustrative purposes only. The fact that in the Household Budget Surveys there is no information about tastes, poses limitations in the examination of the role of tastes versus resource constraints as determinants of the availability of various housing amenities and consumer durables (for example a household may have the ability to buy a car but decided not to do so). Further, there is no information about the quality of the stock of durable goods available to each household, thus, possibly blurring the differences in the living standards of various members of the population (for example, better-off households are likely to have more expensive cars than less well-off households).

The estimate of the Spearman rank correlation coefficient between permanent income and INDEX1 is 0.450 against estimates of 0.409-0.426 in the case of the other monetary welfare indicators, while the corresponding estimates are 0.562 against 0.509-0.523 in the case of INDEX2 and 0.579 against 0.523-0.536 in the case of INDEX3.

However imperfect the non-monetary indices used may be, the results of Tables 4 and, especially, 5 seem to suggest that the new variable is able to depict better the relative welfare position of the members of the population than the other monetary indicators of welfare employed in the paper.

4. Structure of inequality and poverty

As noted earlier, the level of inequality recorded by the distribution of permanent income is substantially lower than the levels recorded by the distributions of disposable income and consumption expenditure. The next question to be investigated is whether the structures of inequality and poverty as accounted using the distribution of the new welfare indicator differs in significant ways from the corresponding structures as accounted by the other monetary welfare indicators available in the Household Budget Survey. In order to examine the structure of inequality we rely on the mean logarithmic deviation, N

$$N = \frac{1}{n}\sum_{i=1}^{n}\ln\left(\frac{\mu}{y_i}\right) \tag{13}$$

where n is the size of the population, y_i the welfare indicator of person i (income, consumption expenditure, permanent income) and μ the mean of the distribution of this indicator. N is strictly additively decomposable. Thus, if the population is grouped into J mutually exclusive and exhaustive groups, N can be written in the following way that allows the quantification of the contributions of disparities "within" and "between" population groups to aggregate inequality [Anand (1983, Appendix C)]

$$N = \sum_{j=1}^{J}\left(\frac{n_j}{n}\right)N_j + \sum_{j=1}^{J}\left(\frac{n_j}{n}\right)\ln\left(\frac{\mu}{\mu_j}\right) \tag{14}$$

where the subscripts j denote the values of the corresponding variables in group j. The first component in the right hand side of (14) is the contribution of disparities

"within groups" to aggregate inequality – that is, the level of inequality that would have been recorded if the mean of each group's welfare indicator became equal to the aggregate mean by equiproportionate changes in the welfare indicators of the members of the group – while the second term is the "between groups" component of inequality – that is, the level of inequality that would have been recorded if the welfare indicators of the members of each group became equal to the group mean but differences between group means remained intact.

For the purposes of our analysis, the population is grouped into mutually exclusive and exhaustive groups using four alternative criteria: locality, household type, socio-economic group and educational level of the household head. Further, multi-variate decomposition of inequality by population sub-groups is attempted by combining these factors. The proportionate contributions of "between groups" and "within groups" disparities to aggregate inequality according to each of the three welfare indicators for each grouping of the population are presented in Table 6. Ceteris paribus, the larger the number of groups and the more homogeneous the groups, the higher the proportion of aggregate inequality that is attributed to "between-groups" disparities.

The results of Table 6 suggest that the structure of inequality is not affected dramatically by the distribution used. In most population groupings, the "between groups" component of inequality is higher when the distribution of permanent income is used. This is most profound in the case of the contribution of the "between-educational-groups" component. As the evidence of Table 7 (which is discussed below) partly shows, the increase in the proportional contribution of "between-groups" disparities when the distribution of permanent income is used instead of the other distributions occurs despite the fact that in the distribution of permanent income the differences in the group means are not as large as in the distributions of either disposable income or consumption expenditure. The increase in the relative importance of the "between groups" component should be attributed to the fact that our formulation of permanent income mitigates extreme values of current income or consumption expenditure and, thus, influences the level of inequality within particular groups substantially more than it affects the relationship between the group means and, hence, the "between groups" component of inequality. In other words, even though both "between groups" and "within groups" inequalities decline in absolute terms when we move from the distribution of disposable income or consumption

Table 6. Inequality decomposition using alternative concepts of resources

Grouping factor	Number of groups	Contribution of "between groups" inequalities (%)			Contribution of "within groups" inequalities (%)		
		DI	CE	PI	DI	CE	PI
Size of Locality	4	7.7	6.5	8.8	92.3	93.5	91.2
Household Type	9	4.5	6.4	6.2	95.5	93.6	93.8
Socio-economic group of household head	11	12.6	12.1	14.3	87.4	87.9	83.7
Educational level of household head	5	22.6	21.0	25.5	77.4	79.0	74.5
Multi-variate decomposition	560	32.0	28.8	34.6	68.0	71.2	65.4

expenditure to the distribution of permanent income, the disparities "within groups" decline more significantly. From a substantive point of view, the estimates of Table 6 confirm earlier results that inequality in Greece emanates primarily from disparities "within" rather than "between" population groups [Tsakloglou (1993, 1997)]. Only when the population is grouped – into just five groups – according to the educational level of the household head, a substantial proportion of aggregate inequality (21.0%-25.5%) can be attributed to disparities "between groups".

For the purposes of the examination of the structure of poverty under alternative concepts of resources we employ the additively decomposable index of Foster, Greer and Thorbecke (1984), F

$$F = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{z - x_i}{z}\right)^{a} \qquad (15)$$

where z is the poverty line, while x_i representes the "truncated distribution" of the corresponding variable; x_i is equal to y_i when the population member falls below the poverty line and equal to z when the population member lies above it. α is a "poverty aversion" parameter whose value, in line with most empirical studies in the field, is set at $\alpha = 2$, at which the index has a number of desirable properties (focus, monotonicity, transfer sensitivity). When the population is grouped into J mutually exclusive and exhaustive groups, F can be written in the following way that allows

the quantification of the contribution of particular population groups to aggregate poverty

$$F=\sum_{j=1}^{J}\frac{n_j}{n}F_j \qquad (16).$$

In line with several studies of poverty in the European Union [O'Higgins and Jenkins (1990), ISSAS (1990), Hagenaars et al (1994)], we started by setting the poverty line at 50% of the mean of the corresponding distribution. Since permanent income is far more equally distributed than either disposable income or consumption expenditure, the resulting poverty rates using this type of poverty line differ considerably across distributions: 7.80% in the case of permanent income against 16.88% in the case of disposable income and 15.12% in the case of consumption expenditure. It can be argued that these differences may blur the comparisons of the composition of the "poor" across distributions since they focus on different segments of these distributions. For this reason, we also examine the composition of the bottom 20% of the three distributions.

Table 7 reports the population shares, the mean equivalent disposable income, consumption expenditure and permanent income as well as the percentage contributions to aggregate poverty of seven population groups that were found to be high-poverty-risk in earlier studies [Tsakloglou (1990), Tsakloglou and Panopoulou (1998)]: members of rural households, persons aged over 64 living alone, childless couples with at least one member aged over 64, members of households headed by farmers, members of households headed by unemployed persons, members of households headed by pensioners and members of households headed by persons who did not complete primary education. These contributions were calculated using both types of "poverty lines" and should be compared with the population shares of the corresponding groups. As noted above, in almost all occasions, the mean permanent incomes of the high-poverty-risk groups are closer to the national average than their mean disposable incomes or consumption expenditures. Nevertheless, in all but two cases, when the poverty line is set at 50% of the corresponding mean, the contributions of these groups to aggregate poverty are considerably higher when we use as welfare indicator permanent income instead of disposable income or consumption expenditure. On the contrary, when we examine the composition of the poorest 20% of the population in the last three columns of Table 7, in all but one cases the contributions of these groups to aggregate poverty according to permanent income

Table 7. Selected high-poverty-risk groups and their contribution to aggregate poverty according to alternative welfare indicators

Population Group	Populat. Share (%)	Mean equivalent (Greece: 100.00)			Contribution to aggregate poverty (%) Foster, Greer and Thorbecke index ($\alpha = 2$)					
					Poverty line set at 50% of national mean			Bottom 20% of the distribution		
		DI	CE	PI	DI	CE	PI	DI	CE	PI
Members of rural households	24.9	78.1	81.1	82.5	54.2	47.8	56.4	44.4	39.6	42.8
Persons aged over 64 living alone	3.8	75.8	71.9	77.2	11.5	14.0	18.0	6.9	8.7	7.9
Childless couples (at least one aged over 64)	9.0	83.0	80.2	84.1	19.2	21.1	23.4	15.8	16.7	16.6
Members of households headed by farmers	12.4	78.8	82.2	83.4	18.6	16.4	15.4	20.6	18.2	19.8
Members of households headed by unemployed	2.7	72.3	82.3	81.8	6.0	5.1	4.6	4.7	4.7	5.3
Members of households headed by pensioners	24.2	89.1	86.2	89.5	41.9	45.2	52.5	33.1	36.4	35.6
Members of households headed by low educ. persons	17.7	70.1	71.4	75.1	43.4	49.1	55.2	34.2	38.4	38.1

lie between the corresponding contributions according to disposable income and consumption expenditure.

Finally, Table 8 reports the results of logit analysis of the probability of falling below the poverty line using both types of "poverty lines" in the three distributions. Once again the differences between the three indicators are not very large. Nevertheless, it should be noted that in the last three columns where the results are strictly comparable, the overall predictive power of the model is slightly higher when poverty is analysed using permanent income as welfare indicator. From a substantive point of view, the results of Table 8 indicate that residence in rural areas, unemployment and low educational qualifications increase significantly the probability of falling below the poverty line irrespective of the poverty line and the distribution used in the analysis, while other factors such as old age do not always retain their independent statistical significance. At the other end, irrespective of the distribution used, the probability of poverty declines significantly as a result of high educational qualifications, particular occupational characteristics of the household head (employer, non-manual employee or professional self-employed) and, to a lesser extent, residence in big cities.

5. Conclusions

The great majority of empirical distributional studies utilise cross-sectional data on disposable income or, to a lesser extent, consumption expenditure from Income or Budget Surveys. For a number of reasons, such as short interview periods and extensive use of recall questions, in many cases these variables exhibit a lot of artificially high variation. In addition, due to life-cycle factors, in many surveys containing information on both variables they do not exhibit a particularly high degree of correlation. As a result, at least one, and possibly both, cannot be considered as very reliable indicators of the long-run welfare of the members of the population and their use for the design of policies aimed to alleviate poverty, or reduce inequality, may be problematic. The problem is likely to be particularly serious in many developing countries with high levels of poverty where such surveys are conducted at irregular intervals, usually many years apart from one another. In these cases mistakes in the identification of the truly high-poverty-risk groups may have serious consequences in terms of human suffering.

Table 8. Logit parameter estimates of the probability of being in poverty using alternative welfare indicators

Variable	Poverty line set at 50% of national mean			Bottom 20% of the distribution		
	DI	CE	PI	DI	CE	PI
Constant term	-1.645 **	-1.798 **	-3.127 **	-1.382 **	-1.366 **	-1.230 **
Locality						
Cities with population over 100.000	-0.331 **	-0.222 *	-0.377 *	-0.313 **	-0.218 *	-0.292 **
Semi-urban areas	0.581 **	0.635 **	0.771 **	0.491 **	0.566 **	0.573 **
Rural areas	0.749 **	0.386 **	0.805 **	0.671 **	0.393 **	0.577 **
Household type						
One person aged below 65	-0.580 *	-0.525	-0.204	-0.606 *	-0.336	-0.637 **
One person aged 65 or more	0.051	0.784 **	0.888 **	0.064	0.726 **	0.362
Childless couple (both below 65)	-0.594 **	-0.244	-0.266	-0.671 **	-0.237	-0.466 **
Childless couple (at least one over 65)	0.127	0.640 **	0.726 **	0.175	0.524 **	0.353 *
Couple with one child below 18	0.077	-0.214	0.418	-0.058	-0.216	-0.290
Couple with three or more children below 18	0.920 **	0.336	1.108 **	0.786 **	0.385 *	0.521 **
Mono-parental household	0.807 **	-0.077	0.760	0.821 **	0.001	0.203
Other household types	-0.274 *	0.137	0.122	-0.282 *	0.065	-0.089
Socio-economic group of HH head						
Employer in non-agriculture	-0.805 **	-1.362 **	-2.233 **	-0.899 **	-1.781 **	-1.538 **
Professional self-employed in non-agriculture	-0.487 *	-0.544 **	-0.745 *	-0.343 *	-0.351 *	-0.792 **
Non-professional self-employed in non-agric.	0.006	-0.279	-0.310	0.065	-0.346 *	-0.629 **
Farmer or agricultural worker	-0.013	-0.176	0.025	0.151	-0.131	-0.158
Non-manual employee in non-agriculture	-1.062 **	-0.638 **	-0.710 *	-0.853 **	-0.711 **	-0.891 **
Unemployed	1.122 **	0.857 **	1.368 **	1.032 **	0.874 **	1.077 **
Pensioner	0.322 *	0.080	0.451 *	0.344 **	0.109	0.094
Other	0.422 **	0.172	0.585 **	0.468 **	0.091	0.160
Educational level of HH head						
Tertiary education completed	-2.408 **	-1.712 **	-1.956 **	-2.214 **	-1.754 **	-2.070 **
Upper secondary education completed	-0.759 **	-1.043 **	-1.094 **	-0.744 **	-0.990 **	-0.968 **
Lower secondary education completed	-0.390 *	-0.652 **	-0.629 *	-0.356 **	-0.509 **	-0.534 **
Primary education not completed	0.653 **	0.755 **	0.713 **	0.566 **	0.744 **	0.743 **
Overall prediction	83.64%	85.23%	92.24%	80.96%	80.92%	81.60%

* significant at the 5% level (Wald χ^2 test)

** significant at the 1% level (Wald χ^2 test)

The present paper presented a simple methodology, that can be easily replicated in other data sets, for extracting information about a more stable long-run welfare indicator of the population members ("permanent income") under assumptions that cannot be considered as particularly restrictive. The resulting indicator is related to all the available monetary welfare indicators, with the corresponding weights determined endogenously and being inversely related to the degree of "noisiness" of the transitory component of each monetary welfare indicator.

Then, an application was provided using the data of a Greek Household Budget Survey. As anticipated, the distribution of permanent income was found to exhibit substantially lower inequality than the distributions of either disposable income or consumption expenditure. Moreover, permanent income was found to be more closely correlated than the other two monetary welfare indicators to a number of non-monetary welfare indicators that were constructed using the information available in the Household Budget Survey. The structure of inequality, as accounted by the three welfare indicators, did not differ substantially across distributions, although in the distribution of permanent income differences "between groups" were found to account for a slightly higher proportion of aggregate inequality than in the distributions of disposable income and consumption expenditure. Likewise, but depending on the poverty line used, the contributions of a number of high-poverty-risk groups to aggregate poverty were found to be larger using the distribution of permanent income than either of the other distributions. These findings have obvious implications for the design of policies aimed to reduce aggregate inequality and, particularly, for the purposes of targeting efficiently the limited resources available for poverty alleviation.

REFERENCES

Abul Naga R. (1994) “Identifying the poor: A multiple indicator approach”, London School of Economics, Distributional Analysis Research Programme, Discussion Paper No 9.

Abul Naga R. and Burgess R. (1997) “Prediction and determination of household permanent income”, London School of Economics, Distributional Analysis Research Programme, Discussion Paper No 32.

Anand S. (1983) *Inequality and poverty in Malaysia: Measurement and decomposition*, Oxford University Press, New York etc.

Anand S. and Harris C. (1994) "Choosing a welfare indicator”, *American Economic Review (Papers and Proceedings)* 84, pp 226-231.

Bartholomew D. (1984) "The foundations of factor analysis ", *Biometrika* 71, pp 221-32.

Becker G.S. (1965) “A theory of the allocation of time”, *Economic Journal* 75, pp 493-517.

Chaudhuri S. and Ravallion M. (1994) "How well do static indicators identify the chronically poor?", *Journal of Public Economics* 53, pp 367-394.

Cowell F.A. (1995) *Measuring inequality*, (2nd ed.), Prentice Hall / Harvester Wheatsheaf.

Dimelis S., Kollintzas T. and Christodoulakis N.M. (1997) *Economic fluctuations and growth in Greece and Europe*, Stamoulis, Athens (in Greek).

Falk B. and Lee B. (1990) "Time-series implications of Friedman’s permanent income hypothesis", *Journal of Monetary Economics* 26, pp 267-283.

Foster J.E., Greer J. and Thorbecke E. (1984) "A class of decomposable poverty measures", *Econometrica* 52, pp 761-766.

Garfinkel I. and Haveman R.H. (1977) *Earning capacity, poverty and inequality*, Academic Press, New York.

Greene W.H. (1993) *Econometric analysis* (2nd ed.), New York, MacMillan.

Hagenaars, A.J.M., de Vos, K. and Zaidi, M.A. (1994) *Poverty statistics in the late 1980s: Research based on micro-data*, Theme 3, Series C, Eurostat Luxembourg.

Hall R. (1978) "Stochastic implications of the life cycle-permanent income hypothesis: Theory and evidence", *Journal of Political Economy* 86, pp 971-987.

Hodrick R.J. and Prescott E.C. (1997) “Postwar U.S. business cycles: An empirical investigation”, *Journal of Money Credit and Banking* 29, pp 1-16.

ISSAS (1990) *Poverty in figures: Europe in the early 1980s*, Theme 3, Series C, Eurostat, Luxembourg.

McGregor P.P.L. and Borooah V.K. (1992) “Is low spending or low income a better indicator of whether or not a household is poor: Some results from the 1985 Family Expenditure Survey”, *Journal of Social Policy* 21, pp 53-69.

Mercader-Prats M. (1998) “Identifying low standards of living: Evidence from Spain”, in Slottje D.J. (ed.) *Research on Economic Inequality*, Vol. 8, JAI Press, Greenwich.

O'Higgins M. and Jenkins S., 1990, "Poverty in EC: Estimates for 1975, 1980 and 1985", in R. Teekens and B.M.S. van Praag (eds) *Analysing poverty in the European Community*, Eurostat News Special Edition, Luxembourg.

Patterson K. (1992) “The service flow from consumption goods with an application to Friedman’s permanent income hypothesis”, *Oxford Economic Papers* 44, pp 289-305.

Sen A.K. (1992) *Inequality reexamined*, Clarendon Press, Oxford.

Tsakloglou P. (1990) "Aspects of poverty in Greece", *Review of Income and Wealth* 36, 1990, pp 381-402.

Tsakloglou P. (1993) "Aspects of inequality in Greece: Measurement, decomposition and inter-temporal change: 1974, 1982", *Journal of Development Economics* 40, pp 53-74.

Tsakloglou P. (1997) "Changes in inequality in Greece in the 1970s and the 1980s", in Gottschalk P., Gustafsson B. and Palmer E. (eds.) *Changing patterns in the distribution of economic welfare: What happened during the 1980?*, Cambridge University Press, Cambridge.

Tsakloglou P. and Panopoulou G. (1998) “Who are the poor in Greece? Analysing poverty under alternative concepts of resources and equivalence scales” *Journal of European Social Policy* 8, pp 229-252.