DE ECONOMIST 139, NR. 3, 1991

SUBJECTIVE POVERTY LINE DEFINITIONS

BY

ROBERT J. FLIK AND BERNARD M.S. VAN PRAAG*

1 INTRODUCTION

Despite the current slight improvement of the worldwide economic situation, research on *poverty* is still of great importance. Although poverty was thought to be a vanishing phenomenon in Western society in the seventies, in the eighties it has been generally recognized that poverty is still a problem. Given this intuitive feeling, voiced by political parties and other pressure groups, it appears to be rather difficult to identify who is poor. The easiest characterization is by means of current income y_c. A threshold value y_p^* is defined, which is called the *poverty line*, and an individual is identified as poor if his current income $y_c < y_p^*$. One of the drawbacks of this method is that we know of cases where people have sufficient income but are nevertheless restricted to such an extent that they do not feel well-off at all. It may be that such people have bad health, suffer from a lack of democratic freedom, or that their income can only be earned by making extremely long working weeks. We may also think of the Russian example (1990) where people frequently have enough money, but there are no commodities in the shops to be purchased. All these examples point to the fact that in many cases a state of poverty has to be described by more than one criterion, in this case income. Although a lot of attention is paid to this multi-dimensional aspect, the poverty concept is invariably operationalized by means of only one characteristic, *i.e.* income, and we shall conform to this tradition. The reasons for this choice are the difficulties of operationalizing multi-dimensional concepts.

In the past, many methods have been developed with a view to constructing a poverty line in terms of income. Those methods may be divided into two groups: the *absolute* measurement methods and the *relative* measurement methods. In the first category, the poverty lines are determined regardless of the general welfare level and the welfare distribution in the society concerned.

* Research Assistant, Econometric Institute, Erasmus University Rotterdam, and Professor of Economics, Econometric Institute, Erasmus University Rotterdam and member of the Scientific Council for Government Policy (W.R.R.), The Hague.

In the second category, poverty is seen as a purely relative phenomenon. People are considered poor if their financial conditions are below the average level of wealth in society. Thus, according to these methods, poverty is considered as a phenomenon of inequality. In the absolute approach it is possible that nobody is poor in a specific society according to the absolute definition. In the relative approach there will always be an 'underclass', the members of which are poor by definition.

As a first attempt to identify poverty in an *absolute* way we mention the basic needs approach. In this method a certain minimum cost diet is determined, which secures physical survival of an individual. The amount of money that one needs in order to buy this food parcel is considered to represent the poverty line. Rowntree (1901) suggested to add the costs of other indispensable needs to survive to the minimum cost diet such as clothes and housing. Orshansky (1965) proposed to multiply the poverty line based on the above mentioned food parcel with the reciprocal of the average Engel coefficient *i.e.* the reciprocal of the average food-income ratio (c/y). This factor was estimated to equal about 3. This is still the basis of the U.S. poverty concept. An obvious deficiency in these methods is the rather arbitrary specification of the minimum cost diet.

An alternative method is based on Engel's observation (1883) that the food-income ratio monotonically decreases when household incomes increase. The food-income ratio is then a proxy for welfare (see *e.g.* Teekens and Zaidi (1989)). A specific level of the food-income ratio is taken to be the poverty threshold; families with an actual food-income ratio higher than this threshold are considered to be *poor* and vice versa. The definition of 'food' is a problem. Does it include luxuries or not? A second objection to this method is that there are huge differences in 'food expenditure attitudes' among people. For instance, there may be rich people who spend the greater part of their income on all kinds of extraordinarily expensive components within the food parcel. Such people would be called poor according to this method, which seems rather strange. An advantage of the Engel ratio method is that it 'automatically' adjusts for differences in household size. A larger family will need a larger amount of food than a small family. Hence, a large family is more likely to live in poverty at a fixed income than a small family.

An example of a *relative* poverty concept (*e.g.* Townsend (1974)) is the definition of a certain fixed percentile, say 25%, of the income distribution as the poverty borderline. But the problem then is that 'the poor are always with us.' A way to circumvent this drawback is to define the poverty line as a certain percentage of the average (or median) income in society (*e.g.* Fuchs (1967)). Then it may happen that nobody will be defined as poor. A general drawback of relative poverty concepts is that the amount of poor people remains unchanged if all incomes increase or decrease by the same percentage.

Comparing the absolute and relative concepts it must be observed that the absolute notions are less absolute than they seem at first. For the absolute criteria (*e.g.* food expenditure or Engel ratio) are fixed in such a way that they are socially acceptable. They have the tendency to be somewhat higher in wealthier societies.

In this paper we will focus on poverty line definitions, which are not *a priori* meant to be either absolute or relative. They take as point of departure the *perception* of poverty as viewed by the individual members of society. More specifically, it is assumed that individuals themselves are the best judges of their own situation. The resulting poverty thresholds will be called 'subjective poverty lines.' Hagenaars and Van Praag (1985) have shown empirically that subjective poverty lines in some sense may be seen as a mixture of absolute and relative concepts.

First we will discuss the *Leyden Poverty Line* (*LPL*), called after its place of origine, introduced in Goedhart *et al.* (1977). This method is based on the *Welfare Function of Income* (*WFI*) $U(y)$, which is derived from a particular survey question, the *Income Evaluation Question* (*IEQ*). The exact wording of this question as well as the derivation of the *WFI* can be found in section 2. According to the *LPL*, which will be the topic of section 3, families are called poor if their after-tax family income falls below an income amount, which corresponds to a specific utility/welfare[1] level δ measured by the *WFI*.

Although the *LPL* is a subjective poverty line as well, Kapteyn, Van de Geer and Van de Stadt (1985) reserved the term *Subjective Poverty Line* (*SPL*) for a specific variant which we shall consider in section 4. This was also introduced by Goedhart *et al.* (1977). This method states that families are poor if their incomes are not sufficient 'to make ends meet' according to their opinion. This measure is based on a one-level attitude question: the *Minimum Income Question* (*MINQ*). It may be seen as a simplified version of the *IEQ*.[2]

The *Centre for Social Policy Poverty Line* (*CSP*) introduced by Deleeck (1977) and later on improved by Deleeck *et al.* (1984) will be described in section 5 (see also Deleeck and Van den Bosch (1989)). Deleeck based this method on the *MINQ* and on an attitude question, very similar to the *IEQ*. We will see that this method only uses a subsample of the respondents who filled in the questionnaire. An alternative method will also be given in this section, which will be more or less based on the ideas of Deleeck, but which will not *a priori* dismiss people from the survey.

Some empirical results of the three subjective poverty line definitions will be presented in section 6. We will also consider the *reliability* for each measure, measured by its standard deviation. This reliability assessment is important if we want to get accurate estimates of the number of people living in poverty. It will be seen that the *LPL* performs best in this respect, followed by the *SPL*, while *CSP* ranks third. Finally section 7 concludes.

1 We will use the terms utility and welfare indiscriminately.

2 See *e.g.* Colasanto *et al.* (1984) and Danziger *et al.* (1984) for other empirical studies concerning the *SPL*.

2 THE INDIVIDUAL WELFARE FUNCTION OF INCOME (WFI)

When deriving a subjective poverty line, one requires, on average, that people who are called poor according to the corresponding measurement method should indeed consider themselves poor, while people with household incomes above the poverty threshold should indeed consider themselves non-poor. So, we are interested in the way each individual thinks about his own financial circumstances and consequently about his own state of well-being. Then the determination of the poverty line is actually based on those welfare judgments.

One way to elicit the individual's welfare judgments is to ask for them by means of survey questions. A particular question which has been posed to thousands of respondents over the years is the so-called *Income Evaluation Question* (*IEQ*), derived by Van Praag (1971). As we will see, the *Leyden Poverty Line* (*LPL*) is fully determined by the outcomes of the *IEQ*, combined with little additional knowledge of personal characteristics. The *IEQ* goes as follows:

> 'Please try to indicate what you consider to be an appropriate amount for your household for each of the following cases. Under my/our conditions I would call an after-tax household income per week/month/year of:
> about ..... very bad,
> about ..... bad,
> about ..... insufficient,
> about ..... sufficient,
> about ..... good,
> about ..... very good.
> Please enter an answer on each line and underline the period you refer to.'

We denote the responses to this question of individual i by $c_{i1}, \ldots, c_{i6}$ and $(i = 1, \ldots, n)$ respectively. Here we state six verbal labels, whereas in other questionnaires fewer or more levels have been presented.

If we also know the incomes of the respondents we are able to examine the way in which each individual evaluates his own (household) income. Besides this we also have information about how the respondents evaluate other income levels. As a next step, we specify a relationship between after-tax household income (on a continuous scale) and corresponding 'numerical valuations of well-being' (also on a continuous scale). Obviously, this requires a transformation of the verbal labels like 'good' and 'bad' into a numerical scale. The resulting function U will be called the (*cardinal*) utility function of income or the *Welfare Function of Income* (*WFI*).[3]

In economic literature there has been a lot of discussion on the question whether utility can be measured, and even on the question whether it exists as a meaningful concept (Robbins (1932)). Some people stated that cardinal utili-

3 In Van Praag and Van der Sar (1988) the *IEQ* has been used for an ordinal analysis.

ty is an immeasurable concept altogether; other people like Pareto (1909) for instance did not state that cardinal utility was an immeasurable concept, but only that it was not necessary to measure utility in order to set up an operational theory of demand behaviour. According to that view, knowledge of the *ordinal* utility function and the corresponding contour lines (*i.e.* indifference curves) is all we need to explain demand behaviour of consumers. See also Van Praag (1968, 1988, 1989) for more elaborate discussions on this issue.

Sen (1986) stresses the existence of two approaches in economics in his booklet on *Economics and Ethics* (1986). On the one hand, there is the 'engineering' approach which tries to model *behaviour*. On the other hand, there is the 'ethical' approach which is basic to the *evaluation* of distribution and allocation. According to Sen, the latter approach is inadequate. It needs cardinal concepts. With ordinal information alone we cannot compare utility values between individuals, because in the 'ordinal way' the utility function is not unique. When we like to make interpersonal comparisons we explicitly need uniquely defined interpersonally comparable cardinal utility functions. This is also an almost necessary first step in the construction of meaningful social welfare functions $W(U_1, \ldots, U_n)$ where social welfare is a function of individual utilities. We shall not pursue this idea any further in this paper.

In Van Praag (1968) a theoretical framework is developed suggesting that the *WFI* can approximately be described by a lognormal distribution function with *individual* parameters μ_{1i} and σ_{1i}^2. In Van Praag (1971) and Van Praag and Kapteyn (1973) estimates on the *WFI* have been given for the first time. See also Kapteyn, Wansbeek and Buyze (1980) and Hagenaars (1986). In Van Praag (1989) new empirical evidence on the lognormality of the *WFI* is given which was based on the outcomes of an attitude question in which people were asked to give evaluations in terms of numbers and line segments of varying length corresponding to different verbal labels, not referring to any subject matter in general or income in particular. For individual i we specify his *WFI* as:

$$U_i(y) = \Lambda(y; \mu_{1i}, \sigma_{1i}^2) = N(ln(y); \mu_{1i}, \sigma_{1i}^2) \tag{2.1}$$

where $\Lambda(\ldots)$ and $N(\ldots)$ denote the lognormal and normal distribution functions respectively.

The parameters μ_{1i} and σ_{1i}^2, corresponding to individual i, can be derived from the answers of the *IEQ* in the following way[4]:

$$\mu_{1i} = \tfrac{1}{6} \sum_{j=1}^{6} ln(c_{ij}) \tag{2.2}$$

$$\sigma_{1i}^2 = \tfrac{1}{5} \sum_{j=1}^{6} (ln(c_{ij}) - \mu_{1i})^2 \tag{2.3}$$

4 The method to estimate the individual μ_{1i} and σ_{1i}^2 differs from the traditional method mentioned in Van Praag (1971). See Van Praag (1989).

Thus, when we have gathered the individual responses to the *IEQ* we have enough information to determine the individual *WFI*'s.

Before we return to the poverty concept, let us consider the μ_1 parameter of the *WFI*.[5] If μ_1 increases by $\Delta\mu_1$, it implies that the individual i needs $exp(\Delta\mu_1)y_i$ more than before to reach the welfare level, previously corresponding to income level y_i. Thus, μ_1 can be interpreted as a *need parameter*. It has been shown that the variation of this parameter among individuals may be explained by personal characteristics. There is overwhelming evidence (see *op. cit.*) that μ_1 can, to a large extent, be explained by the equation:

$$\mu_1 = \beta_0 + \beta_1 \cdot ln(y_c) + \beta_2 \cdot ln(fs) \tag{2.4}$$

where y_c denotes the after-tax current household income and fs the individual's family size. Adding an $N(0, \sigma^2)$ distributed error term ε to this equation, the parameters can be estimated by Ordinary Least Squares (OLS) or Weighted Least Squares (WLS).[6]

Obviously, we expect the estimate of β_2 to be positive, because an increase of family size will have a cost-increasing effect ('needs rise'); additional expenses have to be incurred, which means that extra income is needed to attain the same welfare as before. Moreover, we also suspect a positive estimate of β_1 which reflects a well-known phenomenon, described in psychophysical Adaptation Theory (Helson (1964)); it is said there that individuals adapt their judgments concerning certain phenomena to their own circumstances. In the case of income evaluation this tendency has been called *preference drift* (see Van Praag (1971)).

3 THE LEYDEN POVERTY LINE (LPL)

We now turn to the definition of poverty. According to the *LPL* a family is called *α-poor* if the evaluation of total family income falls below a certain level of utility α, say, 0.4 or 0.5 (*LPL*(04) and *LPL*(05) respectively). If, on the contrary, the evaluation of the family income exceeds the level α the family is called non-poor. Thus, the *individual* poverty line $y_{\alpha i}$ is defined by solving: $U_i(y) = \alpha$, *i.e.*:

$$\Phi\left(\frac{ln(y) - \mu_{1i}}{\sigma_{1i}}\right) = \alpha \tag{3.1}$$

where $\Phi(.)$ denotes the cumulative distribution function of the standard normal distribution. The solution yields:

5 The utility (on a zero-one scale) derived from an income amount of $exp(\mu_1)$ equals 0.5.

6 To make the sample more representative of the whole population, weights can be determined for each individual. The resulting regression technique is called Weighted Least Squares.

$$ln(y_{\alpha i}) = \mu_{1i} + \sigma_{1i} \cdot \Phi^{-1}(\alpha) \qquad (3.2)$$

Let us now take account of the structural relation (2.4). We substitute (2.4) in (3.2) and omit the individual indices. This implies:

$$ln(y_\alpha) = \beta_0 + \beta_1 \cdot ln(y_c) + \beta_2 \cdot ln(fs) + \sigma_1 \cdot \Phi^{-1}(\alpha) \qquad (3.3)$$

Fixing $ln(fs)$ at a certain amount and σ_1[7] at the population average $\bar{\sigma}_1$, the individual $LPL(\alpha)$ poverty line is drawn as a function of $ln(y_c)$ in Figure 3.1. If it is assumed that α represents the utility threshold value of poverty, we notice that people with family log-income below $ln(y_\alpha^*)$ would classify themselves as *α-poor* because their individual poverty line falls above their own income; likewise, people with family log-income above $ln(y_\alpha^*)$ will call themselves non-poor. Consequently, $ln(y_\alpha^*)$ can be seen as the *national* log-poverty line and can be computed by setting $ln(y_\alpha) = ln(y_c)$ which implies:

$$ln(y_\alpha^*) = \frac{\beta_0 + \beta_2 \cdot ln(fs) + \bar{\sigma}_1 \cdot \Phi^{-1}(\alpha)}{1 - \beta_1} \qquad (3.4)$$

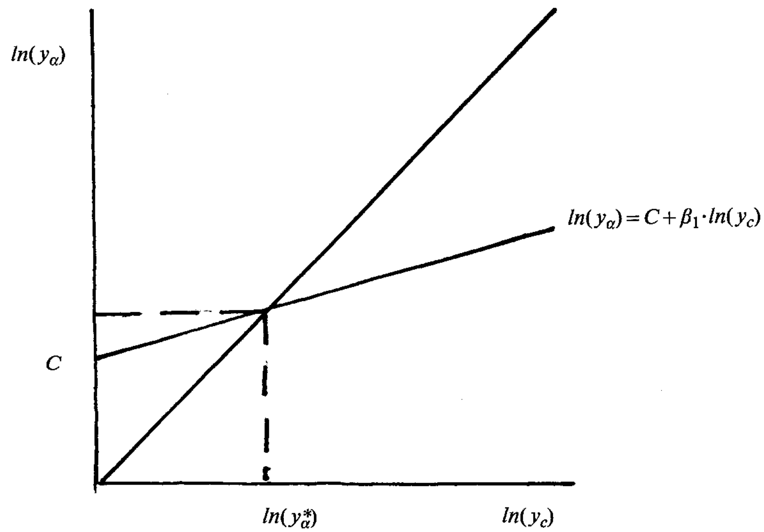

Figure 3.1 – Derivation of the $LPL(\alpha)$ poverty line

7 It has been found that σ_1 depends on $ln(y_c)$ and $ln(fs)$ only marginally; we shall follow the current traditional approach and keep this welfare sensitivity parameter constant.

Thus, we can derive for each family size[8] fs an income amount ($y_\alpha^*(fs)$) which will be the borderline between the poor and the non-poor ('a family-size differentiated poverty line'). For 'national' purposes we can either assess the poor in various groups differentiated according to fs or replace $ln(fs)$ by its average $\overline{ln(fs)}$. In this expository paper we shall do the latter. In a similar way it is conceivable to define poverty lines for subgroups, differentiated according to other variables like age, area, status, health, climate, *etc.* (see *e.g.* Van Praag (1988)).

If the poverty line has to be utilized for politically relevant statements, *e.g.*, for assessing the number of people in poverty or for setting the statutory minimum income level such that a household is eligible for social assistance when it earns less than that minimum income level, it is of prime importance to have an idea of the reliability (or accuracy) of the poverty line. Obviously, $ln(y_\alpha^*)$ is a random outcome, as β_0, β_1 and β_2 are random estimates themselves.[9] To judge the accuracy of this poverty line definition relative to other poverty line definitions we need an assessment of the *variance* of $ln(y_\alpha^*)$ which may be approximated by means of the δ-method (Rao (1973)) as:

$$\sigma^2(ln(y_\alpha^*)) = g^{\mathrm{T}} \cdot C \cdot g \tag{3.5}$$

where

$g^{\mathrm{T}} = [\partial ln(y_\alpha^*)/\partial\beta_0,\ \partial ln(y_\alpha^*)/\beta_1,\ \partial ln(y_\alpha^*)/\partial\beta_2]$ and

C = matrix of variances/covariances of the β-estimators in the μ_1 regression (2.4).

Using this variance and assuming approximate normality of $ln(y_\alpha^*)$ it can be said that with 95% probability the 'true' national family-differentiated (log-) poverty line lies in the range:

$$(ln(y_\alpha^*) - 2 \cdot \sigma(ln(y_\alpha^*)),\ ln(y_\alpha^*) + 2 \cdot \sigma(ln(y_\alpha^*))).$$

Another way in which the variance or the standard deviation may be assessed is by the computer-intensive *bootstrap method* (*e.g.* Efron (1982)). We will use the latter method in this paper, as the δ-method is not applicable on the *CSP*-poverty line definition.

Now that the (Leyden) poverty line has been defined, it is interesting to know the amount of people or the percentage of people in a certain country, *the*

8 As it appears that μ_1 can be explained by variables other than family size and household income, like education, age, *etc.*, the poverty line can also be differentiated with respect to those variables. For the sake of convenience and for ease of exposition, however, we will only take family size into account.

9 We neglect the randomness of σ_1.

poverty ratio, who are to be called poor according to this measure. This poverty ratio may be computed by simple counting. In order to get a structural idea about the relation between the poverty ratio and the income distribution it is worthwhile to approximate the income distribution by a simple functional specification. Just like the *WFI* the income distribution function can be approximately described by a lognormal distribution function with parameters μ_0 and σ_0^2 (Aitchison and Brown (1957)). As a result the poverty ratio is equal to:

$$\int_0^{y_\alpha^*} \mathrm{d}\Lambda(y;\mu_0,\sigma_0^2) = \int_{-\infty}^{ln(y_\alpha^*)} \mathrm{d}N(ln(y);\mu_0,\sigma_0^2) = \Phi\left(\frac{ln(y_\alpha^*)-\mu_0}{\sigma_0}\right) \tag{3.6}$$

Finally, we like to know the average log-income and the average income of the people who are called poor according to this measurement method.[10] To compute these amounts, we need the mathematical expressions of the first moment of the truncated standard normal distribution function given in (3.6) and the first moment of the truncated lognormal distribution function (Johnson and Kotz (1971)). The average log-income equals:

$$\int_{-\infty}^{ln(y_\alpha^*)} ln(y)\,\mathrm{d}N(ln(y);\mu_0,\sigma_0^2) = \frac{-\varphi\left(\dfrac{ln(y_\alpha^*)-\mu_0}{\sigma_0}\right)}{\Phi\left(\dfrac{ln(y_\alpha^*)-\mu_0}{\sigma_0}\right)}\cdot\sigma_0+\mu_0 \tag{3.7}$$

and average income equals:

$$\int_0^{y_\alpha^*} y\,\mathrm{d}\Lambda(y;\mu_0,\sigma_0^2) = exp(\mu_0+\tfrac{1}{2}\sigma_0^2)\cdot\frac{\Phi\left(\dfrac{ln(y_\alpha^*)-\mu_0}{\sigma_0}-\sigma_0\right)}{\Phi\left(\dfrac{ln(y_\alpha^*)-\mu_0}{\sigma_0}\right)}. \tag{3.8}$$

Again $\Phi(\cdots)$ denotes the cumulative distribution function of the standard normal distribution, whereas $\varphi(\cdots)$ denotes its density function.

4 THE SUBJECTIVE POVERTY LINE (SPL)

Instead of asking income amounts, which correspond to *several* welfare levels, it is also possible to ask for *only one* income amount, which corresponds to a specific welfare label, which is assumed to describe the boundary between

10 Of course there are more and even better measures, which describe the extent of poverty in different countries, but in this paper we will restrict ourselves to the percentage of poor people and the average income of these people. See for more details *e.g.* Hagenaars (1986).

'poor' and 'non-poor'. This particular survey question is called the *Minimum Income Question* (*MINQ*) and goes as follows:

> 'What do you consider as an absolute minimum net income for a household such as yours? In other words, we would like to know an income amount below which you won't be able *to make both ends meet.*
> about ..... per week / per month / per year
> Please underline the period you refer to.'

Let the answer of individual i be $c_{min,i}$. According to the *Subjective Poverty Line* definition (*SPL*) (see *e.g.* Kapteyn, Van de Geer and Van de Stadt (1985); see also the original definition in Goedhart *et al.* (1977)) this answer is called the individual's poverty line. Obviously, $c_{min,i}$ depends on personal characteristics, just as the answers of the *IEQ* (or as μ_{1i} which is a 'compound measure' representing these *IEQ* answers: see section 2). Again, we take into account the own current household income and the family size of the respondent. We specify the following log-linear relationship:

$$ln(c_{min}) = \gamma_0 + \gamma_1 \cdot ln(y_c) + \gamma_2 \cdot ln(fs) \tag{4.1}$$

Adding a $N(0, \sigma^2)$ distributed error term ε to this equation, the parameters can be estimated by OLS or WLS. For the same reasons as given in section 3 we suspect the estimates of both γ_1 and γ_2 to be positive.

For various sizes of the family we draw the individual (log-) poverty lines as functions of $ln(y_c)$ in Figure 4.1.

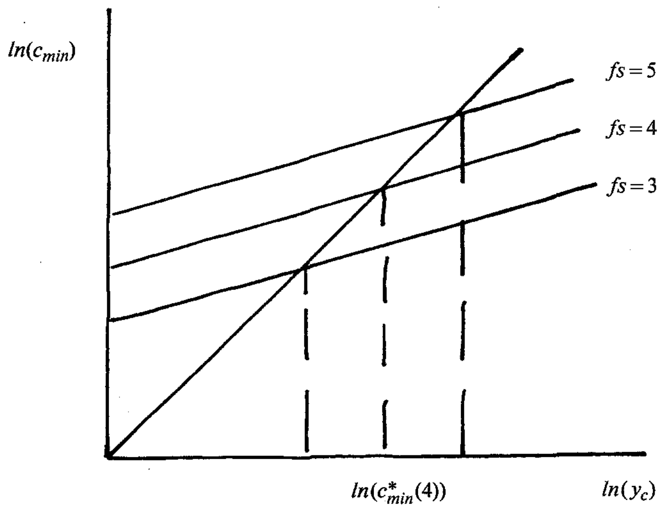

Figure 4.1 – Derivation of the *SPL* poverty line for varying household sizes

Now, we like to derive family-size differentiated *national* poverty lines. By setting $ln(y_c)$ equal to $ln(c_{min})$ for each family size we obtain these national log-poverty lines $ln(c^*_{min}(fs))$ which equal:

$$ln(c^*_{min}(fs)) = \frac{\gamma_0 + \gamma_2 \cdot ln(fs)}{1 - \gamma_1} \tag{4.2}$$

This is evident in Figure 4.1; families with four persons, for instance, will call themselves poor if their log-income falls below $ln(c^*_{min}(4))$; on the other hand, if they have a log-income which lies above $ln(c^*_{min}(4))$ they will classify themselves as non-poor.

Finally, analogous to the *LPL*, we may derive the standard deviation of the *SPL* by means of the bootstrap method, and the percentage of poor people and the average income of these people by substituting $ln(c^*_{min})$ (or $ln(c^*_{min}(fs))$) for $ln(y^*_\alpha)$ in equations (3.6) and (3.8) respectively.

The *LPL* (based on a multi-level question) seems to be theoretically superior to the *SPL* (based on a one-level question), as the *SPL* is likely to be more subject to random response fluctuations and more sensitive to varying interpretations of the one level. The answers to several ordered verbal labels may be expected to be much more carefully selected and calibrated as one has to rank several levels than when the respondent is offered just one level, *i.e.*, 'make ends meet.' With one stimulus only there is more room for interrespondent variation in interpretation. Moreover, in the multi-level case it is possible to interpolate between the verbal levels given in the questionnaire and to calculate measures like μ_{1i} and σ^2_{1i} by which the random response errors at the separate levels will cancel each other out to a large extent. This possibility is absent when we use the *SPL*. Finally, the *LPL* definition is also preferable from a political view, because politicians may choose a specific utility level (between zero and one), and the poverty line belonging to that utility level with the corresponding characteristics may be computed. On the other hand, when they use the *SPL* definition they have to accept the outcomes on a take-it-or-leave-it basis. Results for a European data set have been given in Van Praag, Goedhart and Kapteyn (1980), in Van Praag, Hagenaars and Van Weeren (1982) and in an extensive analysis by Hagenaars (1986). See also Danziger *et al.* (1984) and Colasanto *et al.* (1984) for studies on U.S. data.

5 THE CENTRE FOR SOCIAL POLICY POVERTY LINE (CSP)

At *The Centre for Social Policy* (Antwerp), Professor Deleeck and his co-workers independently developed the *CSP*-measure. It uses the *MINQ*, some standard information, and a special multi-level attitude question which we, for convenience, call the *'Deleeck question'* and which reads as follows:

'Can you make ends meet with the actual net income of your household:
- with great difficulty,
- with difficulty,
- with some difficulty,
- rather easily,
- easily,
- very easily?'

The respondent has to mark only one category, *viz.*, the response category which his income fits best; Deleeck then singles out the subsample of respondents who classified their own income under the heading 'with some difficulty.' For each remaining respondent in the 'with some difficulty' group the minimum between the current household income y_c and the registered minimum income c_{min} (from the *MINQ*) is defined as the respondent's 'lower income' (*ylow*).

Next, the average and the standard deviation of these *ylow* values are calculated. After rejection of the outliers for which *ylow* differs by more than two standard deviations from the average *ylow*, a new average for *ylow* is computed, to be used as the poverty threshold. The standard deviation of this measure cannot be obtained in a simple analytical way, but may again be computed by using the bootstrap method.

The procedure can be performed for the whole sample, but also for different subgroups differentiated according to sex, age and/or family size, for instance.

If in a particular questionnaire the Deleeck question is not available, and the *IEQ* has been posed instead, we can apply the same method using the *IEQ*. We single out those respondents whose actual household income lies in the 'insufficient' range $\{\frac{1}{2} \cdot (ln(c_2) + ln(c_3)), \frac{1}{2} \cdot (ln(c_3) + ln(c_4))\}$, which is assumed to correspond with the 'with some difficulty' range used in the original Deleeck wording. The same procedures can also be followed using the *MINQ*. Then, after determination of the poverty line the familiar characteristics can be computed in the usual way.

As we can see, Deleeck imposes a strong assumption; he argues that the level of the poverty line must be fixed by people who are on the margin of poverty and consequently have first-hand knowledge of the situation. At this point we see a considerable difference between the *LPL* and the *SPL* on the one hand and the *CSP* on the other. The *CSP* method uses a small subsample of people considering themselves on the margin, while in the *LPL* and *SPL* procedures the opinions on what constitutes a situation of poverty and non-poverty of *both* people who consider their own income on the margin *and* of those who consider their own income in a different way (*below* or *above* the margin) are utilized for the estimation of the poverty threshold. Even if we admit that people who earn far more than the poverty line will have a hazy notion of what is 'with some difficulty,' it is not democratic to discard the opinion of the richer and especially the poorer part of the population from the information set. The whole population determines the social norm. Besides, a result of this screening procedure of the *CSP* is the strong dependence on the choice of the sample of respondents. If there are only a few people who belong to the reference group, only a few people determine the poverty line and the random factor becomes a strong influence.

An *alternative method* which is more or less based on Deleeck's ideas, however without using the *MINQ*, and which is not excluding respondents is the following method. We assume that we have outcomes of an *IEQ*, again with answers $c_{i1}, \ldots, c_{i6}$ $(i = 1, \ldots, n)$. Then to start with this method, we estimate the following equations:

$$\begin{aligned} ln(c_2) &= \beta_{02} + \beta_{12} \cdot ln(y) + \beta_{22} \cdot ln(fs) + \varepsilon_2 \\ ln(c_3) &= \beta_{03} + \beta_{13} \cdot ln(y) + \beta_{23} \cdot ln(fs) + \varepsilon_3 \\ ln(c_4) &= \beta_{04} + \beta_{14} \cdot ln(y) + \beta_{24} \cdot ln(fs) + \varepsilon_4 \end{aligned} \tag{5.1}$$

Instead of estimating these equations one by one using Ordinary Least Squares or Weighted Least Squares these equations may also be estimated simultaneously using a (weighted) 'Seemingly Unrelated Regression' method (SUR).[11] If we would apply OLS to the three separate equations we would obtain the same estimators as with applying SUR, because the three equations contain the same explanatory variables; however, the advantage of using SUR is the assessment of the matrix of variances/covariances of the β-estimators, which we would need when we apply the δ-method to compute the variance of the poverty line.

The *CSP* determines the poverty threshold by computing the average of the 'lower incomes' of the people who are in the 'with some difficulty' area, which was constructed by that reference group. We think, however, that a poverty line must be determined by taking into account the opinions of both these people and all the other people who filled in the *IEQ*. So in contrast to the *CSP*, we prefer computing the average income in the 'with some difficulty' area, which is constructed by *all* people in the survey. This 'insufficient' range equals:

$$\{\tfrac{1}{2} \cdot (ln(c_2) + ln(c_3)), \quad \tfrac{1}{2} \cdot (ln(c_3) + ln(c_4))\} \tag{5.2}$$

11 Suppose we would like to estimate the model $Y_i = X_i\beta_i + e_i$ $i = 1, \ldots, M$. If one then focuses attention on one equation, say the ith, the Ordinary Least Squares estimator $\beta_i = (X_i'X_i)^{-1}X_i'Y_i$ is the minimum variance, linear unbiased estimator. However, we can improve on this estimator by taking into account the correlation between e_i and the other disturbance vectors. Consequently, a better estimate is the ith component of $\beta = (X'(\Sigma^{-1} \otimes I)X)^{-1}X'(\Sigma^{-1} \otimes I)Y$, where $\otimes$ denotes the Kronecker product and Σ the $M*M$ covariance matrix of the errors where $E(e_ie_j') = \sigma_{ij} \cdot I$.

or after substituting (5.1) in (5.2) and omitting the error terms:

$$\begin{aligned}&\{\tfrac{1}{2}\cdot[(\beta_{02}+\beta_{03})+(\beta_{12}+\beta_{13})\cdot ln(y)+(\beta_{22}+\beta_{23})ln(fs)],\\ &\tfrac{1}{2}\cdot[(\beta_{03}+\beta_{04})+(\beta_{13}+\beta_{14})\cdot ln(y)+(\beta_{23}+\beta_{24})ln(fs)]\}\end{aligned} \tag{5.3}$$

For each log-income $ln(y)$ which lies between these boundaries the following holds:

$$\begin{aligned}&\frac{\tfrac{1}{2}\cdot[(\beta_{02}+\beta_{03})+(\beta_{22}+\beta_{23})\cdot ln(fs)]}{1-\tfrac{1}{2}\cdot(\beta_{12}+\beta_{13})}\leq ln(y)\leq\\ &\frac{\tfrac{1}{2}\cdot[(\beta_{03}+\beta_{04})+(\beta_{23}+\beta_{24})\cdot ln(fs)]}{1-\tfrac{1}{2}\cdot(\beta_{13}+\beta_{14})}\end{aligned} \tag{5.4}$$

or

$$A\leq ln(y)\leq B \tag{5.5}$$

where A and B are shorthand notations for the lower and upper bounds in (5.4).

If we would like to compute the average $ln(y)$ between these boundaries for different family sizes, or for the population as a whole by fixing $ln(fs)$ at the average $\overline{ln(fs)}$ and if we again assume lognormally distributed family incomes, we need the first moment of the *double-truncated* standard normal distribution. It can be shown that this average log-income equals (see Johnson and Kotz (1971)):

$$In\ (y^*)=\frac{\varphi\left(\frac{A-\mu_0}{\sigma_0}\right)-\varphi\left(\frac{B-\mu_0}{\sigma_0}\right)}{\Phi\left(\frac{B-\mu_0}{\sigma_0}\right)-\Phi\left(\frac{A-\mu_0}{\sigma_0}\right)}\cdot\sigma_0+\mu_0 \tag{5.6}$$

Consequently, this amount can be considered as the log-poverty line. The poverty line y^* may be computed as follows: we know from (5.5):

$$exp(A)\leq y\leq exp(B) \tag{5.7}$$

To compute the average income between these boundaries, we need the first moment of the *double-truncated* lognormal distribution. This average income equals (again see Johnson and Kotz (1971)):

$$y^*=exp(\mu_0+\tfrac{1}{2}\sigma_0^2)\cdot\frac{\Phi\left(\frac{B-\mu_0}{\sigma_0}-\sigma_0\right)-\Phi\left(\frac{A-\mu_0}{\sigma_0}-\sigma_0\right)}{\Phi\left(\frac{B-\mu_0}{\sigma_0}\right)-\Phi\left(\frac{A-\mu_0}{\sigma_0}\right)} \tag{5.8}$$

Finally, the standard deviation, the poverty ratio and the average income of the poor people can be computed in the usual way, again assuming family incomes to be lognormally distributed with parameters μ_0 and σ_0^2.

In the next section we will discuss some empirical results for the reviewed poverty line definitions.

6 SOME EMPIRICAL RESULTS

In this section we present and briefly discuss some empirical results of the application of the poverty line definitions mentioned above. To this end we used the answers of a Dutch survey, conducted in 1983 by The Netherlands Joint Press Office (GPD). In this questionnaire, along with a list of individual and household characteristics, the people were asked to fill in both the *Income Evaluation Question* and *The Minimum Income Question*. With the aid of the so-called RAS method (see Hagenaars (1986)) every individual in the sample has been given a weight, which makes the sample representative for the population.

Let us first give the results of some estimated relationships, which will be used in the computation of the poverty thresholds and the corresponding characteristics.

For the $LPL(\alpha)$ poverty line(s) we need the outcomes of equation (2.4). For each individual μ_{1i} can be calculated according to equation (2.2). Then OLS regression yields[12]:

$$\begin{array}{llll} \mu_1 = 2.6763 + 0.7112 \cdot ln(y_c) + 0.0841 \cdot ln(fs) & R^2 = 0.6270 \\ \quad\;\;(0.0738)\;\;(0.0070) \qquad\qquad (0.0063) & N\;\;= 6313 \end{array} \tag{6.1}$$

where y_c denotes the after-tax household income and fs the family size.

For the SPL poverty line we estimated the $ln(c_{min})$ relation (4.1) as follows:

$$\begin{array}{llll} ln(c_{min}) = 3.8447 + 0.5855 \cdot ln(y_c) + 0.1088 \cdot ln(fs) & R^2 = 0.3050 \\ \qquad\qquad\;(0.1207)\;\;(0.0115) \qquad\qquad (0.0103) & N\;\;= 6313 \end{array} \tag{6.2}$$

Looking at these relations, we conclude that the SPL method must be more sensitive to response fluctuations than the $LPL(\alpha)$ methods, because of the better statistical performance of regression (6.1) (higher R^2 and lower standard deviations).

We also estimated the equations (5.1). The results are:

12 The figures in parentheses are standard deviations.

$$\begin{aligned}
ln(c_2) &= \underset{(0.0882)}{3.1132} + \underset{(0.0084)}{0.6440} \cdot ln(y_c) + \underset{(0.0075)}{0.1137} \cdot ln(fs) \quad R^2 = 0.5055 \\
ln(c_3) &= \underset{(0.0826)}{2.5879} + \underset{(0.0078)}{0.7068} \cdot ln(y_c) + \underset{(0.0070)}{0.1083} \cdot ln(fs) \quad R^2 = 0.5779 \qquad (6.3) \\
ln(c_4) &= \underset{(0.0788)}{2.2310} + \underset{(0.0075)}{0.7596} \cdot ln(y_c) + \underset{(0.0067)}{0.0902} \cdot ln(fs) \quad R^2 = 0.6199
\end{aligned}$$

Next, we will give some (weighted) average values of variables in the sample and some constants, which we will also need. The average value of (log-)family size equals 1.1336, whereas the average welfare sensitivity parameter of the lognormal welfare function of income ($\bar{\sigma}_1$) equals 0.3135. Furthermore, we need $\Phi^{-1}(\alpha)$ values for different α levels. In this section we only take into account $\alpha = 0.4$ and $\alpha = 0.5$; their corresponding $\Phi^{-1}(\alpha)$ values are -0.25335 and 0.00 respectively. Finally, with the aid of a statistical table, constructed by the Dutch Central Statistical Office (CBS), we compute the parameters of the approximately lognormal distribution of Dutch after-tax household incomes (1983). We found: $\mu_0 = 10.3490$ and $\sigma_0^2 = 0.3577$.

Then as a result the national, fixed-family-size poverty thresholds, the corresponding standard deviations, the theoretical percentages of poor people and the average poor incomes according to the five different methods are tabulated in Table 6.1. We also computed the percentage of people who earn an income which lies in the 2σ-probability interval around the poverty line to see for how many people it is questionable whether they must be called poor or non-poor.

Of course, it is also possible to compute *family size differentiated* poverty

TABLE 6.1 - POVERTY LINES, STANDARD DEVIATIONS, PERCENTAGES, AVERAGE INCOMES AND POPULATION SHARES IN THE 2σ-PROBABILITY INTERVAL

	pov.line•	st.dev••	pov.ratio[13]	avg.inc.•	2σ-pop.share
LPL(04)	11182	555	4.3%	8950	3.0%
LPL(05)	14722	562	10.4%	11373	4.6%
SPL	15411	604	11.9%	11824	5.2%
Original *CSP*	25572	778	34.4%	17239	7.7%
Alternative *CSP*	11307	601	4.4%	9013	3.3%

• amount in Dfl.

•• computed by means of the bootstrap method; 200 samples of 6313 observations have been generated.

13 Of course these expected percentages belong to the average poverty lines as given in the first colum of this table. However, taking into account the standard deviation of these poverty lines, the 'true' percentages do not have to equal these percentages, but will lie in a certain range built around the expected percentages.

lines and consequently family equivalence scales; however, it is not our aim to go into details here (see for more detailed results applied to more data sets *e.g.* Hagenaars (1986)). Looking at this table, we immediately see the high threshold amount according to the original *CSP* method. This is in line with the recent results of comparative studies by Ghiatis (1989) and Moriani (1989). Furthermore, it is remarkable that the *LPL*(04) and the 'alternative *CSP*' ended up with nearly the same poverty line and corresponding characteristics. The reason for this conformity may be the fact that each individual, probably unconsciously, divides the income-axis into six intervals, whereby each succeeding interval is given an additional 16.67% of 'total utility.'[14] Then it turns out that the 40% utility level corresponds with approximately the middle of the third interval, which happens to be the 'with some difficulty' range.

At the same time we notice the conformity between the *LPL*(05) results and the *SPL* results. From this we might conclude that the income amounts, which were filled in as the answers of the *Minimum Income Question*, were evaluated at about 0.5 on the zero-one utility scale (on average).

Finally, if we compare the standard deviations of the five different poverty lines, as a measure of the instability of the threshold values, we conclude that the original *CSP* method has the highest amount of uncertainty about the 'true' value of this poverty threshold. Moreover, the small but nevertheless present difference in the standard deviations of the *LPL*(α) lines and the *SPL* line indicates that the *SPL* is slightly more subject to random fluctuations, but of course this higher standard deviation is an indirect consequence of the higher standard deviations in the c_{min} regression.

7 DISCUSSION AND CONCLUSION

In this paper we outlined three different methods, which are supposed to identify poverty in a subjective way in the sense that *people were asked* to express their feelings about it. This is done by conducting surveys, in which people are asked to answer, among other things, certain attitude questions in which they have to give value judgments in terms of income amounts corresponding to welfare labels (*IEQ*, *MINQ*) and/or have to insert their own household incomes in one of several welfare-labeled income brackets (*Deleeck question*).

We described the way each method derives a poverty line in terms of after-tax household income, which separates the poor from the non-poor. We found that the *SPL* is slightly more subject to random response fluctuations than the *LPL*. We think that people answer more accurately if they have to fill in several levels (*IEQ*) instead of giving just one value judgment (*MINQ*). On the other hand, the Deleeck method (*CSP*) is based on the assumption that poverty threshold values must be determined by the opinions of only those people who

14 This phenomenon is often called the *Equal Quantile Assumption* (*EQA*) (see *e.g.* Van Praag (1989)).

are actually on the margin of poverty. We prefer, however, to take into account the opinions of *all* respondents. For that reason we proposed an alternative method which is more or less based on the ideas of Deleeck, but which does not dismiss respondents from the analysis.

At the end of this paper estimation results were given for a Dutch data set created in 1983. The above-mentioned predictions about the *SPL* and the original *CSP* were empirically verified. Besides, it turned out, among other things, that the results of the *LPL*(04) poverty line and the alternative *CSP* method were quite similar; this is probably due to an assumption concerning the way the 'average respondent' fills in the *IEQ*, *i.e.* the *Equal Quantile Assumption* (*EQA*), which was often taken for granted in earlier research programmes.

Finally, we should note that more extensive surveys, which are more representative of the national population, are needed to yield outcomes of the different poverty lines, which should be fully comparable.

REFERENCES

Aitchison, J. and J.A.C. Brown (1957), *The Lognormal Distribution*, Cambridge.

Colasanto, D., A. Kapteyn and J. van der Gaag (1984), 'Two Subjective Definitions of Poverty: Results from the Wisconsin Basic Needs Study,' *Journal of Human Resources*, 19, pp. 127–137.

Danziger, S., J. van der Gaag, M.K. Taussig and E. Smolensky (1984), 'The Direct Measurement of Welfare Levels; How Much Does it Cost to Make Ends Meet?,' *Review of Economics and Statistics*, 66, pp. 500–505.

Deleeck, H. (1977), 'Bestaanszekerheid en het Sociale Zekerheidsstelsel in België–1974,' *Belgisch Tijdschrift voor Sociale Zekerheid*, 19, pp. 1–23.

Deleeck, H., J. Berghman and P. Janssens (1984), 'Sociale Zekerheid: Inkomensverdeling en Doelmatigheid Anno 1982,' *De Gids Op Maatschappelijk Gebied*, 75, pp. 595–623.

Deleeck, H. and K. Van den Bosch (1989), 'The Measurement of Poverty in a Comparative Context: Empirical Evidence and Methodological Evaluation of Four Poverty Lines in Seven EC-countries,' Report presented at Noordwijk, Seminar *Poverty Statistics in The European Community*.

Efron, B. (1982), *The Jackknife, the Bootstrap and Other Resampling Plans*, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania.

Engel, E. (1883), 'Der Werth des Menschen. 1. Teil: Der Kostenwerth des Menschen,' *Volkswirtschaftliche Zeitfragen*, 37-38, pp. 1–74.

Fuchs, V. (1967), 'Redefining Poverty and Redistributing Income,' *Public Interest*, Summer, pp. 88–95.

Ghiatis, A. (1989), 'Low Income Groups Obtained by Enhanced Processing of the Household Budget Surveys in the EC – Summary Figures for Italy and The Netherlands,' Report presented at Noordwijk, Seminar *Poverty Statistics in the European Community*.

Goedhart, Th., V. Halberstadt, A. Kapteyn and B.M.S. van Praag (1977), 'The Poverty Line: Concept and Measurement,' *The Journal of Human Resources*, 12, pp. 503–520.

Hagenaars, A.J.M. and B.M.S. van Praag (1985), 'A Synthesis of Poverty Line Definitions,' *Review of Income and Wealth*, 31, pp. 139–153.

Hagenaars, A.J.M. (1986), *The Perception of Poverty*, Amsterdam.

Helson, H. (1964), *Adaptation-Level Theory; An Experimental and Systematic Approach to Behaviour*, New York.

Johnson, N.L. and S. Kotz (1971), *Distributions in Statistics, Continuous Univariate Distributions-1*, New York.

Kapteyn, A., T.J. Wansbeek and J. Buyze (1980), 'The Dynamics of Preference Formation,' *Journal of Economic Behaviour and Organization*, 1, pp. 123–157.

Kapteyn, A., S. van de Geer and H. van de Stadt (1985), 'The Impact of Changes in Income and Family Composition on Subjective Measures of Well-Being,' in: M. David and T. Smeeding (eds.), *Horizontal Equity, Uncertainty and Economic Well-Being*, Chicago.

Moriani, C. (1989), 'The Italian Experience on Poverty Harmonized Measures,' Report presented at Noordwijk, Seminar *Poverty Statistics in the European Community*.

Orshansky, M. (1965), 'Counting the Poor: Another Look at the Poverty Profile,' *Social Security Bulletin*, 28, pp. 3–29.

Pareto, V. (1909), *Manuel d'économie politique*, Paris.

Praag, B.M.S. van (1968), *Individual Welfare Functions and Consumer Behavior*, Amsterdam.

Praag, B.M.S. van (1971), 'The Welfare Function of Income in Belgium: An Empirical Investigation,' *European Economic Review*, 2, pp. 337–369.

Praag, B.M.S. van and A. Kapteyn (1973), 'Further Evidence on the Individual Welfare Function of Income: An Empirical Investigation in The Netherlands,' *European Economic Review*, 4, pp. 33–62.

Praag, B.M.S. van, Th. Goedhart and A. Kapteyn (1980), 'The Poverty Line. A Pilot Survey in Europe,' *Review of Economics and Statistics*, 62, pp. 461–465.

Praag, B.M.S. van (1982), 'A Comparison of Objective and Subjective Measures of Poverty,' in: G. Sarpellon (ed.), *Understanding Poverty*, Istituto Internazionale 'Jacques Maritain,' Roma, Italy and Franco Angeli Libri s.r.l. Milano, Italy.

Praag, B.M.S. van, A.J.M. Hagenaars and J. van Weeren (1982), 'Poverty in Europe,' *Review of Income and Wealth*, 28, pp. 345–359.

Praag, B.M.S. van (1988), 'The Relativity of the Welfare Concept,' Report WP 69, World Institute for Development Economics Research of the United Nations University, Helsinki, forthcoming in: M. Nussbaum and A. Sen (eds.), *The Quality of Life*, Oxford.

Praag, B.M.S. van and N.L. van der Sar (1988), 'Household Cost Functions and Equivalence Scales,' *Journal of Human Resources*, 23, pp. 193–210.

Praag, B.M.S. van (1989), 'Ordinal and Cardinal Utility: An Integration of the Two Dimensions of the Welfare Concept,' forthcoming in *The Journal of Econometrics*.

Rao, C.R. (1973), *Linear Statistical Inference and its Applications*, New York.

Robbins, L. (1932), *An Essay on the Nature and Significance of Economic Science*, 1st. ed., London.

Rowntree, B.S. (1901), *Poverty: A Study of Town Life*, London.

Sen, A. (1986), *Economics and Ethics*, London.

Teekens, R. and M.A. Zaidi (1989), 'Relative and Absolute Poverty in The European Community; Results from Family Budget Surveys,' Report presented at Noordwijk, Seminar *Poverty Statistics in the European Community*.

Townsend, P. (1974), 'Poverty as Relative Deprivation: Resources and Style of Living,' in: D. Wedderburn (ed.), *Poverty, Inequality and Class Structure*, London.

Summary

SUBJECTIVE POVERTY LINE DEFINITIONS

In this paper we will deal with definitions of *subjective* poverty lines. To measure a poverty threshold value in terms of household income, which separates the poor from the non-poor, we take into account the opinions of all people in society. Three subjective methods will be discussed and compared, *viz.*, the Leyden Poverty Line (*LPL*), the Subjective Poverty Line (*SPL*) and the Centre for Social Policy Poverty Line (*CSP*). In the empirical part of the paper we compute the average poverty line and a few corresponding characteristics for each definition.