Applied Economics Letters, 2002, **9**, 69–72

Do we tend to overestimate poverty gaps? The impact of equivalency scales on the calculation of the poverty gap

BRANKO MILANOVIC

World Bank, Development Economics Research Department, Washington DC, USA
E-mail: bmilanovic@worldbank.org

The standard formulae used to calculate poverty gaps are incorrect when the poverty analysis is conducted in terms of equivalent units (rather than in per capita terms). An exact calculation requires access to individual level data. If this is unavailable, the paper proposes an approximation.

I. DEFINING THE PROBLEM: AN ILLUSTRATION

Consider a population that consists of only two households. The first household has a total household income of 500 and is composed of three people: 2 adults and a child. Let, for simplicity, the equivalence scale be such that each adult counts for 1, and each child for 0.5. That the household's income per consumption unit[1] is 200 (500 divided by 2.5 equivalent units). The second household has a total income of 450 and is composed of 2 people: 1 adult and a child. Its per consumption unit income is 300 (450 divided by 1.5). Let now the poverty line per equivalent unit be 400. The analysis is conducted in terms of consumption units, and a household is considered as poor if its income per consumption unit is less than the poverty line. Both households are poor, and poverty headcount (P_0) is 1.

The question is: What is the amount of the poverty gap? The overall poverty gap is defined as the total amount of money that should be given to the households to bring them up to the level of the poverty line. For the first household, the total 'required' income not to be poor is 1000 (400 times 2.5 equivalent units). For the second household, it is 600 (400 times 1.5). Since the combined income of the two households is 950, and the 'required' amount is 1600, the overall poverty gap is 650.

This study will now show what the application of the standard formulae will yield. Overall poverty gap is equal to:

$$PG = P_1 Nz = P_0 INz \tag{1}$$

where PG = poverty gap, P_1 and P_0 are FGT measures with a respectively 1 and 0, I = average per capita income shortfall (of the poor) as percentage of the poverty line, N = total population (both poor and non-poor) and z = poverty line. From our example, the values of the variables are $P_0 = 1$, $N = 2$, $z = 400$ and thus:

$$I = \frac{1}{Q}\sum_{i=1}^{5}\frac{z - y_i^e}{z} = \frac{1}{5^*400}[3^*(400-200) + 2^*(400-300)] = 0.40 \tag{2}$$

where Q is the total number of the poor individuals (= 5 in this example), and y_i^e = per adult equivalent income of the i-th individual. Each individual is imputed y_i^e of his/her household. $I = 0.4$ means that, on average, income shortfall of the poor as percentage of the poverty line is 40%. If all the values are plugged back into Equation (1) it becomes:

$$PG = 1^*0.4^*5^*400 = 800 \tag{3}$$

and thus an overall poverty gap of 800 which is greater than the poverty gap calculated above (650).

[1] Terms 'equivalent' or 'consumption' unit are used interchangeably.

Applied Economics Letters ISSN 1350–4851 print/ISSN 1466–4291 online © 2002 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/135048501100493 97

Equations 1–3 are the standard formulae used both when the analysis is conducted in per capita terms (individuals are ranked according to their household per capita income) and in per equivalent adult terms (individuals are ranked according to household equivalent income). The only difference in the latter case is that in Equation 2 y_i^e appears instead of y_i (average per capita income of i-th household).

However, as seen, the formulae do not seem to be correct. The mistake is concealed in the 'weighing' in Equation 2 where the difference between the poverty line and household equivalent income ($z - y_i^e$) is multiplied by the number of people in the household rather than by the number of equivalency units in the household.

To see this, this study will go back to its original example where the income shortfall for the first household. The income shortfall for that household was (400–200) times 2.5 consumption units = 500. Or in Equation 2, the income shortfall is calculated as (400–200) times 3 persons. It seems that income shortfall is overestimated by the factor by which the number of persons among the poor exceeds the number of consumption units.

II. THE CORRECT OVERALL POVERTY GAP

The correct income shortfall of i-th household is:

$$I_i = w_i(z - y_i^e) \tag{4}$$

where w_i = number of consumption units in the household. The overall poverty gap then becomes:

$$\sum_{i=1}^{2} I_i = \sum_{i=1}^{2} w_i(z - y_i^e) = 2.5^*(400 - 200) + 1.5^*(400 - 300) = 650 \tag{5}$$

If the overall poverty gap changes, clearly the average poverty gap as percentage of the poverty line must also change. By definition, new I^* will be:

$$I^* = \frac{\sum_{i=1}^{H} I_i}{Qz} = \frac{\sum_{i=1}^{H}\sum w_i(z - y_i^e)}{Qz} = \frac{zW - \bar{y}Q}{Qz} = \frac{W}{Q} - \frac{\bar{y}}{z} = 0.8 - \frac{190}{400} = 0.325 \tag{6}$$

where $i = 1$ to H indicates that the summation is done over households, W = total sum of the consumption units among the poor, and this study uses the fact that total income of the poor is $Sw_i y_i^e = Q\bar{y}$ where $\bar{y}$ = average per capita income of the poor. In other words, whether the sum equivalent household incomes across consumption units (w_i) or per capita household incomes across individuals, the sum is always equal to total income. This study's problem arises when household equivalent incomes are summed across individuals.

How does the correct I^* differ from I? Equation (2) can be written as

$$I = \frac{1}{Q}\sum_{i=1}^{Q} \frac{z - y_i^e}{z} = \frac{Qz - \sum y_i^e}{Qz} \tag{7}$$

Then,

$$\frac{I^*}{I} = \frac{\dfrac{Wz - Q\bar{y}}{Qz}}{\dfrac{Qz - \sum y_i^e}{Qz}} = \frac{Wz - Q\bar{y}}{Qz - \sum y_i^e} \tag{8}$$

It is known that $Sy_i^e > Q\bar{y}$ because the sum of household equivalent incomes across individuals must be greater than total income. Also, $Qz > Wz$. Thus both terms in the denominator of Equation 8 are greater than the corresponding terms in the numerator. What is the percentage extent of the overestimation of the gap? It is equal to $1 - (I^*/I)$ because all the other terms in the poverty gap (P_0, N and z)] are the same. Replacing the values from the example in Equation 8, it is seen that the gap is overestimated by 18.75% (see Equation 9) while, as this study saw above, the number of consumption units among the poor is 20% less than the number of individuals (Equation 4 rather than 5):

$$\frac{4^*400 - 5^*190}{5^*400 - 1200} = \frac{650}{800} = 0.8125 \tag{9}$$

III. CAN THE EXACT OVERESTIMATE BE EASILY CALCULATED?

It is seen that as a first approximation the overestimation of the gap (using the standard formula) is equal to the ratio between the number of persons and number of consumption units among the poor. Thus if it is known, for example, that the average number of people per poor household is 4 and the average number of consumption units 3, it can be guessed that the gap will be approximated by the ratio 4-to-3, or 33%. The issue is now the following: can the extent of the gap overestimation be exactly calculated if only the aggregate numbers are had, i.e. the data on total number of consumption units among the poor and total number of the poor (W and Q) as well as the N, z and $\bar{y}$. These values are often easily available.

Let first the number of consumption units for each household be expressed as $w_i = a_i n_i$, where a_i = specific ratio for that household ($0 < a_i < 1$), and n_i = number of persons in i-th household. Then, Equation 8 becomes:

$$\frac{I^*}{I} = \frac{\sum_i^Q a_i n_i z - Q\bar{y}}{Qz - \sum_{j=1}^{H} \dfrac{Ti}{a_i n_i}}. \tag{10}$$

where $Y - i =$ total household income. Now, if the ratio between the number of consumption units and individuals is constant across the poor, $a_i = a$ and Equation 10 simplifies:

$$\frac{I^*}{I} = \frac{aQz - Q\bar{y}}{Qz - \frac{1}{a}Q\bar{y}} = a$$

The overestimate can be readily calculated only in this very simple case when the equivalent unit/person ratio is constant across the poor households. The lower a, the greater the overestimate of the poverty gap. This is, of course, as expected: if the number of consumption units are small (relative to the number of people) our standard formulae are going to produce significant overestimate of the poverty gap.

But if more realistically, $a_i = f(y)$, i.e. let a_i increase with income per capita since poorer households typically have more children per each adult, the exact overestimation will depend on the pattern or speed with which a_i increase. In general, the faster they increase the less will be the overestimate.[2]

The following example (Table 1) shows the ratio between the standard calculated poverty gap and the actual poverty gap in two countries whose household sizes and compositions are very different. A very simple equivalence scale is used such that the number of consumption units is equal to the square root of the number of individuals living in a household. In Estonia, households are small: the average household size is 2.41, the average number of consumption units 1.49 (ratio 0.62). In Kyrghyzstan, both are, of course, larger: the average household size is 4.93, and the average number of consumption units 2.13 (ratio 0.43). If poor is defined as all individuals whose household equivalent income places them in the bottom decile, it is found that the consumption unit/household size differences are magnified for the poor. The ratios are 0.57 and 0.4 for respectively Estonia and Kyrghyzstan. As discussed above, these values give us the first order approximation of the ratio between the actual poverty gap and the standard poverty gap. When calculations are done, however, it is found that the overestimate is even greater: in Estonia, the correct ratio is 0.54 (and the true poverty gap is overestimated by a factor of 1.86), and in Kyrghyzstan, the ratio is 0.31 (and the true poverty gap is overestimated by a factor of 3.22). But since individual level data is often missing, the first order approximation, shown below, may be the best possible:

$$\text{True poverty gap} = \text{Conventional poverty gap } x\frac{W}{Q}$$

The error implied in the use of the standard formulae is greater the further from the per capita measurement, and since sharper equivalence scale is likely to be used for countries where households are large, the error will tend to be more important in 'younger' countries. Figure 1 shows the correct and the conventional (or 'wrong') P1 calculated using the individual level data from Household surveys from Cote d'Ivoire and Hungary. On the horizontal axis is the adjustment factor for the equivalence scale (?) such that income per equivalent adult is equal to Y/n^{θ}, where ? takes values between 0 and 1. Obviously, if ? = 1, per capita measurement is used, and the lower the ?, the sharper is the adjustment for economies of scale. Thus, ? of 0.5 may be

Table 1. *Household size, consumption units and the poverty gap*

	Estonia	Kyrghyzstan
For the entire population		
(1) Average HH size	2.41	4.93
(2) Average number of consumption units	1.49	2.13
Ratio (1): (2)	0.62	0.43
For the bottom decile (= poor)		
(1) Average HH size	2.95	5.78
(2) Average number of consumption units	1.67	2.33
Ratio (1): (2)	0.57	0.40
Poverty gap (for the bottom decile) a/		
(1) Conventional poverty gap	29963	695578
(2) True poverty gap	16348	213698
Ratio (1): (2)	0.54	0.31

a/ Amounts in local currency.
Source: Calculated from individual survey data from the World Bank HEIDE data base (available at http://www.worldbank.org/research/transition/index.htm).

2

$$\frac{d(I^*/I)}{da}\frac{da}{dy} = \frac{1}{I}\left(z\sum\frac{da}{dy}n_i\right) - \frac{I^*}{I^2}\left(-\frac{1}{a^2}\sum\frac{Y_i}{n_i}\frac{da}{dy}\right) = \frac{1}{I}(zQ)\frac{da}{dy} + \frac{I^*}{I^2}\left(\frac{1}{a^2}Q\bar{y}\right)\frac{da}{dy} = \frac{Q}{I}\left(z + \frac{I^*}{I}\frac{\bar{y}}{a^2}\right)\frac{da}{dy}$$

The ratio I/I^* will increase the greater da/dy > 0, that is the steeper is the increase of a as income per capita goes up.

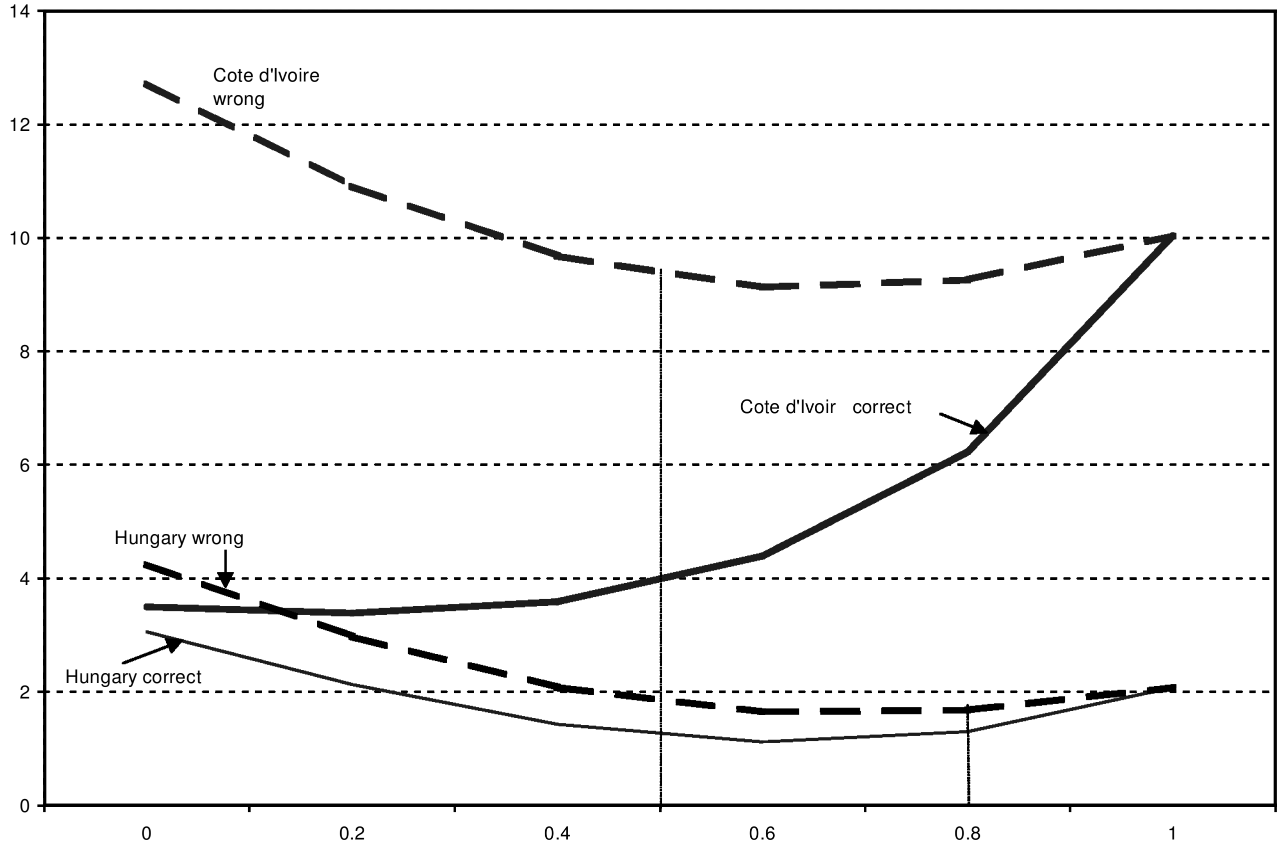

Fig. 1. *The correct and 'wrong' P1 as function of equivalence scale*
Sources: Hungary: Household budget survey 1993; Cote d'Ivoire: Enquete prioritaire sur les dimensions sociales de l'ajustement structurel, 1993. Poverty line is both cases is equal to 50% of mean equivalized expenditures or income.

deemed reasonable for Cote d'Ivoire where the average family size is in excess of six, while a theta of 0.8 may be more appropriate for Hungary where the average household size is less than three. In that case too, the error associated with the calculation of P1 in Cote d'Ivoire it is likely the poverty gap will be significantly overestimated in countries with larger average household size, which also tend to be poorer countries.

IV. CONCLUSIONS

When the poverty analysis is conducted in terms of consumption units (rather than in per capita terms), the conventionally calculated poverty gap will be much larger than the true poverty gap. This will be particularly the case for countries with a large average household size, and therefore low ratio between consumption units and people. To make things worse, these are generally poorer countries (with lower mean income or expenditures) and thus the overestimation of the poverty gap is likely to be particularly large in countries that have most difficulty raising funds to address their poverty. To calculate the true poverty gap one needs access to individual data. Short of that, one can use an approximation: simply multiply the conventionally calculated poverty gap by the ratio between the average number of consumption units and average household size among the poor.

ACKNOWLEDGEMENTS

The author first became aware of this problem when he first worked on world bank poverty assessment for Poland in 1995. The author is grateful to Irena Topinska and Christiaan Grootaert for the discussion of the problem, and helpful comments.