

How robust are equivalent scales?⁽¹⁾

Quanto sono robuste le scale di equivalenza?

Chiara Seghieri

Dip. di Statistica "G. Parenti", V.le Morgagni, 59, 57134 FI, seghieri@ds.unifi.it

Gustavo De Santis

DESMS, V. T. Cannizzaro 298, 98122 ME, gdesantis@unime.it

Riassunto: La costruzione di una scala di equivalenza poggia su un complesso apparato teorico, dai molti punti critici, lungamente dibattuti in letteratura e ancora senza una chiara soluzione. In sordina, però, è rimasto un altro, importante problema: scelto un approccio, cambiare il modo di formare le categorie di spesa incide sensibilmente sul risultato. Appare necessario cercare di capire il perché e superare questo limite perché le scale di equivalenza diventino uno strumento logico e operativo di effettiva utilità.

Keywords: Equivalence scales; complete demand system; estimation; style effects.

1. Introduction

With respect to a reference household r , a household h that differs in one or more respects - say, number and age of members, or presence of disabled persons - can be considered more (or less) "needy", in terms of the economic resources necessary to reach the same material level of wellbeing. Ideally, therefore, a properly constructed equivalence scale ε_h should reveal just how much household h needs, relative to household r , to reach the same level of material wellbeing, and this information is essential in several domains, from taxation to welfare analysis, provision of subsidised goods and services, etc. Unfortunately, equivalence scales also suffer from a few theoretical and empirical shortcomings: they can be constructed in a variety of ways, on the basis of widely differing empirical indicators, relying on alternative, and conflicting, assumptions, etc.; they vary over time and across countries, and no universal consensus has yet been reached on the "best" way to estimate them. As a result, each country uses its own, and several of the equivalence scales currently in use (e.g. OECD, or modified OECD) are based more on policy decisions than on empirical evidence.

Another, relatively neglected, problem with equivalence scales is that, even within a given theoretical approach and for a given time period, results vary markedly depending on an apparently minor detail: the number and type of consumption categories considered. The present paper addresses precisely this problem, by comparing the estimated ε 's obtained with 3 different approaches (Engel, Ray and DM^2), all of which can be applied to a single cross-sectional source of data on household consumption. Theoretically speaking,

¹(□) We retain full responsibility for the processing of Istat's microdata on Household Budgets. We acknowledge financial support from the Italian MIUR and from the University of Messina. Discussion with Mauro Maltagliati and Gianni Marliani (un. of Firenze) proved highly beneficial.

all three are defensible, although the third, DM^2 , is probably preferable (De Santis and Maltagliati, 2003). But how do they perform empirically?

2. Theory

The empirical estimation of an equivalent scale on the basis of consumption data requires a relatively large theoretical background (cf. e.g. Deaton and Muellbauer, 1980). A simpler approach, however, based on "indicators", leads to the same formulae. Let w_g , the consumption share of a generic good g , depend on the equivalent household income $EI=X/\varepsilon$, where X stands for income (or, as in our case, total consumption, used as a proxy). The equivalent scale ε itself depends on the demographic characteristics of the household: for instance, $\varepsilon_h=1+k(N-1)$, where k is the key parameter of interest, and N is the household dimension. Our reference household r has $N_r=1$ and $\varepsilon_r=1$. In practice, a frequent assumption is that, for each household h ,

$$0) \quad w_g = f_g(EI) = f_g\left(\frac{X}{\varepsilon}\right) = f_g\left(\frac{X}{1+k(N-1)}\right)$$

where the function f generally uses the log-income, either linearly or in a quadratic form. Under this assumption, the estimated value of our standard-of-living indicator $\hat{w}_{hg}$ will equal its reference value $\hat{w}_{rg}$ *iff* $EI_h = EI_r$. Since X and N are empirically observable, and $\varepsilon_r=1$ by definition, ε_h , the equivalence scale that we are looking for, can be derived. Now, the general principle underlying eq. (0) can be specified in several different ways. Engel, for instance, applies it exclusively to food, assuming that two households are equally well off when their *food share* is the same

$$1) \quad w_{hf} = a_f + b_f \ln(X_h/\varepsilon_h) + u_{hf} \quad (f=\text{food}; u=\text{random residual}) - \text{ENGEL}$$

Ray (1983) extends Engel's idea: two households are assumed to be equally well off when *all* of their consumption shares (not just food) are the same

$$2) \quad w_{hg} = a_g + b_g \ln(X_h/\varepsilon_h) + u_{hg} \quad (g=1, \dots, G \text{ goods}; \Sigma a_g=1; \Sigma b_g=0) - \text{RAY}$$

De Santis and Maltagliati (2003) extend the notion further: they assume that two households are equally well off when *all* of their consumption shares *tend* to be the same, except for a style effect, that is a modification in the consumption profile (our w 's), independent of equivalent income but dependent on the household characteristic h : for instance, you buy toys if you have a child, you live prevalently indoors if you are old, etc.

$$3) \quad w_{hg} = a_g + b_g \ln(X_h/\varepsilon_h) + s_g[D_h] + u_{hg} \quad (\Sigma s_g=0; \text{rest as before}) - DM^2$$

where, as before, the demographic characteristics of the household can be represented more or less simply: e.g. dimension (N_h-1), or age groups (Young, Adults-1, Old), etc.

Notice that the restrictions on the parameters in equations (2) and (3) limit simultaneous estimation to just $(G-1)$ regression equations, with the G^{th} calculable by difference. Therefore eq. (1) (Engel) can be seen as a particular specification of eq. (2) (Ray), with two consumption items only: food (estimated directly) and non-food (calculated by difference). Similarly, eq. (3) can be specified for two expenditure categories only (for instance, food and, as a residual, non-food), thus obtaining a fourth possibility:

$$4) w_{hf} = a_f + b_f \ln(X_h/\varepsilon_h) + s_f[D_h] + u_{hf} \quad (ENGEL - DM^2)$$

3. Data

For our empirical applications we used the individual (household) data of the Istat Consumption Survey for the years 1997-2003, i.e. post the 1996 revision. Each set, with data from more than 20 thousand households, details almost 280 elementary expenditure categories, which we transformed into shares w and grouped according to three different criteria: w_{istat}) "Economic similarity" between groups: e.g. food, transport, education, etc.; w_{beta}) Relation to log-income. At each step we grouped the two categories with the closest values for b_i , the regression coefficient in eq. 2 and 3. This should create homogeneous expenditure categories, ranging from luxuries to necessities, as b declines; w_{small}) Economic weight. At each step, we grouped the two expenditure categories with the lowest share. This should minimise the number of zeroes in the dependent variable, which is another non-trivial technical problem with household consumption data. We estimate our k (cost of additional member) on 15, 14, ..., 3, 2 groups, for each set.

We keep the demographic part to a minimum, by considering two dimensions only: number (N), and, separately, age group of household members - i.e. we count separately the young Y , the adults A (aged 20-59) and the old O .

All regressions are applied to all years, separately, but here we will present our results only for the most recent year, 2003.

4. Results

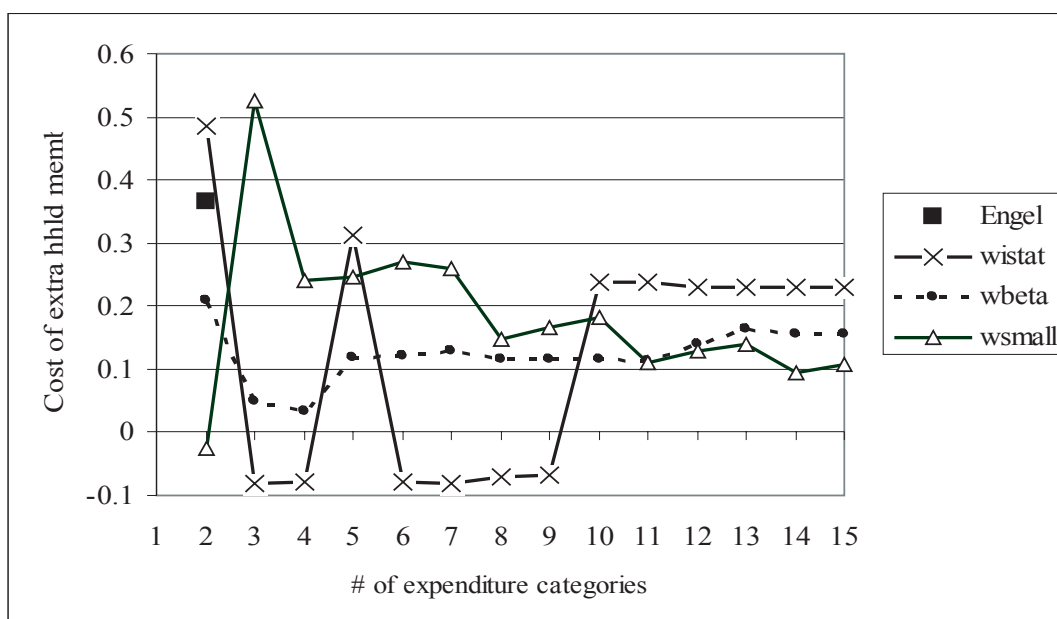
In principle, one should expect to obtain more or less the same equivalence scales (i.e. the same k) when using a given regression method in a given time period, independently of the number and type of consumption categories. Further, equivalence scales should not "jump" too much from one year to the next, or from one functional form to another (eg. using also quadratic forms, so as to improve the fit).

Unfortunately this is not the case. Figure 1, for instance, shows that in 2003, DM^2 estimates of the cost of additional household members change markedly as one varies the composition of the expenditure basket: from negative(!?) to +50% with w_{istat} and w_{small} , and from +3% to +21% with w_{beta} . Similar results (not shown here) emerge with *Ray*, or for other years, of with linear forms.

Our main results are: 1) Engel generally, but not consistently, leads to higher values ($k=37\%$ in figure 1), although they, too, vary considerably with the functional form adopted; 2) grouping on the basis of w_{beta} leads to the smallest variability in the estimates; 3) a large number of expenditure goods (10+) is the best guarantee against the

vagaries of the estimates, although a perfect convergence of results (i.e. independently of the criteria adopted for grouping goods) cannot be obtained; and 4) k (cost of extra member) in 2003 probably lies somewhere between 15% (w_beta) and 25% (w_istat).

Figure 1. *Estimated cost of an extra person in the household, with various criteria and number (2 to 15) of expenditure goods (Italy 2003). Methods: DM^2 and Engel*



5. Conclusions

Equivalence scales are of paramount importance for welfare analysis and policy decisions. Unfortunately, however, they appear scarcely robust to apparently trivial variations in the number and typology of expenditure categories. To our knowledge, this limitation has never been analysed, or simply remarked, before - with reference to either the three models considered here (Engel, Ray, and DM^2) or to any of the several other models frequently discussed in the specialised literature (e.g. scaling, translating, Barten, ...). Notwithstanding this neglect, we consider this a crucial issue, worthy of further investigation, both empirically (as in this paper) and theoretically.

References

- Deaton A., Muellbauer J. (1980) *Economics and consumer behaviour*, Cambridge University Press, Cambridge.
- De Santis G., Maltagliati M., (2003) Complete demand systems: a fresh look at an old problem. Theory and empirical evidence, in *Household behaviour, equivalence scales and well-being*, Dagum C. & Ferrari G. (Eds.), Physica Verlag, 29-53.
- Ray R. (1983) Measuring the cost of children. An alternative approach, *Journal of Public Economics*, 22 (1), 89-102.