THE ESTIMATION OF THE COSTS OF CHILDREN BASED ON RANDOM EQUIVALENCE SCALES

Stanisław Maciej Kot

Gdansk University of Technology, Poland

e-mail: skot@zie.pg.gda.pl

ORCID: 0000-0002-5875-6498

Abstract: This paper measures the cost of children using *the random equivalence scale* (RES). From the general population perspective, the deterministic microeconomic equivalence scales appear as continuous random variables. The stude derived the distribution of RES, assuming the lognormal distribution of income. The truncated distribution of RES can account for possible economies of scale in income or expenditure. The positional measures of the truncated distribution of RES may serve as single equivalence scales. A society's attitude towards inequality may help chose such scales. RES for Poland 2015 was estimated using microdata from the Household Budget Survey. Polish households exhibited a remarkable level of economies of scale in that year and the equivalence scales declined with expenditure. It was observed that the cost of bringing up a child is not constant; generally, it decreases with increasing household size.

Keywords: equivalence scale, lognormal distribution, economies of scale, inequality aversion.

1. Introduction

The aim of this paper was to measure the costs of bringing up children using *the random equivalence scale* (RES) developed herein. RES is a stochastic counterpart of the deterministic equivalence scales offered by the microeconomic theory of consumer behaviour.

Equivalence scales are used to enable assessments of inequality, welfare and poverty in the distribution of income or expenditure (hereafter treated interchangeably). Such appraisals require *homogeneous* populations of comparable units. However, data on income come from

surveys where households are statistical units. The populations of households are *heterogeneous* since households may differ in many aspects other than income, e.g. the size, the composition, the age of the adults, the age of the children, the disabilities of the household members, etc. The adjustment of household's incomes by an equivalence scale gives an *equivalent income* in a *homogeneous* population of *equivalent units* enjoying the same standard of living (Buhmann, Rainwater, Schmaus, and Smeeding, 1988; Jones and O'Donnell, 1995; Ebert and Moyes, 2003).

There are well-known shortcomings of microeconomic equivalence scales. The scales require strong assumptions concerning the relationship between income and needs, and serious identification problems arise in the estimation of equivalence scales (see, e.g. Pollak and Wales, 1979, 1992; Blundell and Lewbel, 1991; Blackorby and Donaldson, 1993; Slesnick, 1998; Chiappori, 2016).

A macroeconomic perspective on microeconomic equivalence scales reveals that in the general population of households, income and expenditure, and their admissible transformations, become *random variables.* As countries' household populations are large, these random variables become continuous, according to the limit theorems.

The study derived the distribution of RES from the distributions of incomes in compared populations of households. A positional measure of RES may serve as a single equivalence scale. The author shows how society's attitude to inequality may help choose such single equivalence scales.

In this paper RESs for Poland in 2015 were estimated assuming the lognormal distribution of expenditure, using statistical microdata from the Polish Household Budget Survey, followed by applying RES to assess the costs of having children.

The rest of this paper has the following structure. Section 2 formulates the theoretical frameworks of RES. In Section 3, the distribution of RES is derived when income obeys the lognormal distribution. Section 4 presents the relationship between society's inequality aversion and the positional measures of RES distribution. Section 5 shows the empirical results of estimating RES. Section 6 concludes.

2. The random equivalence scales

2.1. The standard microeconomic equivalence scales

Let $h = 1, \ldots, m,$ be an indicator of household populations, which differ concerning a particular attribute, e.g. size, demographic composition etc., $h = 1$ was reserved for the reference households consisting of single childless persons, although other specifications are possible. Hereafter, the

term '*h*-household' will denote the household that has the characteristic *h*, while x_h and *y* represent actual incomes or expenditures of *h*-household and the reference household, respectively.

The standard microeconomic equivalence scale for *h*-household, given reference household $h = 1$, is defined as follows:

$$z(\mathbf{p}, u_0, h) = \frac{c(\mathbf{p}, u_0, h)}{c(\mathbf{p}, u_0, 1)}, \quad h = 2, \ldots, m, \tag{1}$$

where $c(\cdot)$ is the household expenditure (cost) function, i.e. the minimum cost of attaining utility level u_0 for a given demographic attribute *h* and prices *p* (Deaton and Muellbauer, 1980). Hereafter, it was assumed that all households face the same prices, thus one may omit the argument *p* in Eq. (1). This does not mean neglecting the impact of prices on equivalence scales, assuming a particular price vector, $\mathbf{p}_o$, faced by all households.

It is useful to express the equivalence scale (1) in terms of expenditure *x*. Inverting the cost function with respect to *u* results in the *indirect utility function* $u = v_h(x)$, $h = 1, \ldots, m$, following the usual assumption that the utility functions are continuous and increasing (see e.g. Donaldson and Pendakur, 2003). The expenditure functions and indirect utility functions are related by identities $y = c(u_0, 1) \leftrightarrow v_1(y) = u_0$ and $x_h = c(u_0, h) \leftrightarrow v_h(x_h) = u_0$, $h = 2, \ldots, m$ (Blackorby and Donaldson, 1993). Then, the equivalence scale (1) will be the ratio of actual expenditure

$$z_h = \begin{cases} \dfrac{x_h}{y}, & \text{for } h = 2, \ldots, m \\ 1, & \text{for } h = 1 \end{cases}, \tag{2}$$

if and only if the households attain the same welfare level u_0, namely if and only if $v_h(x_h) = v_1(y) = u_0$.

Note that equations (1) and (2) do not define a single equivalence scale but *a family of equivalence scales* indexed by *u*. However, demand data do not provide information about *u*. This is the main reason for the non-identification of a single equivalence scale from the demand data alone (Lewbel, 1999, p. 193)[1].

Economists have made various attempts to overcome the non-identification problem. One can identify a single equivalence scale assuming (1) or (2) being independent of *u* (the assumptions called, alternatively, *independence of base (IB)* (Lewbel, 1999) or *equivalence-*

---

[1] In economics, a quantity is said to be identified if it can be calculated or estimated from empirical data.

scale exactness (ESE) (Blackorby and Donaldson, 1993).[2] Several authors have tested these assumptions, but ultimately rejected them (Ebert and Moyes, 2003).

Some authors search for additional sources of identifying information. Subjective data on income or life satisfaction are widely used, either for the indirect estimation of equivalence scales via consumer demand systems or for the direct estimation of equivalence scales. (see e.g. van Praag, 1968, 1991; Kapteyn and van Praag, 1978; Koulovatianos, Schröder, and Schmidt, 2005). This approach's main drawback is that a respondent is meant to assess hypothetical incomes that he/she would never experience.

This drawback is not present in the subjective approach that only uses information on how satisfied a respondent is with his/her *actual income* (e.g. Melenberg and van Soest, 1995; Bellemare, Melenberg, and Van Soest, 2002; Schwarze, 2003; Bollinger, Nicoletti, and Pudney, 2012; Biewen and Juhasz, 2017). The qualifications of the form 'good', 'rather good', 'neither good nor bad', 'rather bad', and 'bad' are the ordinal measurement of household welfare.

Some authors look for objective proxies of welfare, for instance, Jackson (1968) uses the budget share for food, the concept returning to Engel (see Lewbel, 1999, pp. 190-192, for a more exhaustive review of Engel's scales). Recently, Szulc (2009) used several 'wellbeing covariates' for appraising household welfare and estimated equivalence scales for Poland with the matching estimator's help.

2.2. Equivalence scales in the general population of households

Let random variables X_h and Y describe income distributions in the general populations of h-households and reference households, respectively[3]. We assume that X_h and Y are continuous random variables with density functions $f_h(x)$ and $f_y(y)$, and the distribution functions $F_h(x)$ and $F_y(y)$, respectively. In the sequel, we use shorthand $X_h \sim f_h(x)$, $Y \sim f_y(y)$, and $X_h \sim F_h(x)$, $Y \sim F_y(y)$. As countries' household populations are usually large, the limit theorems justify continuity of X_h and Y.

Let indirect utility functions $v_1(y)$ and $v_h(x)$ be continuous and increasing in their arguments (Donaldson and Pendakur, 2003). By applying the well-known formula of transforming continuous random variables (see e.g. Fisz, 1963, p. 39), the following results are obtained.

Result 1. The indirect utility functions are continuous random variables in the general population, namely $V_h = v_h(X_h)$ and $V_1 = v_1(Y)$.

---

[2] See Donaldson and Pendakur (2003, 2006) for more general assumptions.

[3] Capital letters were reserved for random variables and small letters for values of random variables.

Result 1 implies the necessity of a reformulation of 'the if and only if condition' of Equation (2). An essential property of a continuous random variable is *zero probability* that it takes any specific number. Mathematicians refer to such a number as *an almost impossible event*[4]. Thus $P(V_1 = u) = P(V_h = u) = 0$, for all real u. In other words, the fulfilment of the condition in question is *almost impossible* in the general population, therefore utility intervals should be used instead of a single utility level.

Assume a finite number k of disjoint non-empty intervals, which cover the whole domain of V_1 and V_h. Let ith interval (u_i, u_{i+1}), $u_i < u_{i+1}$, be such that $P(u_i < V_h < u_{i+1}) > 0$ and $P(u_i < V_1 < u_{i+1}) > 0$, $i = 1,\ldots, k-1$. Then the condition in question is reformulated as follows:

Definition 1. Two households are equally well-off if and only if the values of their indirect utility functions belong to the same interval of welfare.

Assuming h-households and reference households are equally well-off, in the sense of Definition 1, the following result is obtained:

Result 2. There will be a multitude of ratios (2) which are the realisations of the following positive-valued continuous random variable:

$$Z_h = \frac{X_h}{Y},\qquad(3)$$

with the density function $f_z(z)$ and the distribution function $F_z(z)$, $z > 0$. As X_h and Y are independent, the density function $f_z(z)$ takes the form

$$f_z(z) = \int_0^\infty y f_x(y \cdot z) f_y(y) dy,\qquad(4)$$

(Fisz, 1963, p. 62).

Definition 2. Z_h (3) is the random equivalence scale (RES) for h-households.

The results presented above show how the deterministic micro-economic concept of equivalence scales becomes stochastic if one adopts the 'general population perspective'. To make RES operational and consistent with economic merit, the study introduced some additional notions.

Definition 3. A single random equivalence scale is a positional parameter of RES.

When the RES distribution is known, an analyst might use the mean, the geometric mean, the median or the mode as a single equivalence scale.

---

[4] In the theory of probability P({impossible event}) = 0. However, P(A) = 0 does not necessary imply that *A* is an impossible event.

Section 4 shows how society's inequality aversion can help choose a particular equivalence scale.

Z_h (3) may violate *the economies of scale* in incomes or expenditures since its domain (*z*-values) comprises all positive numbers. The concept of economies of scale is "(…) the mechanism that explains why the cost of living of a family is less than the sum of costs of living of its members taken independently (…)" (Chiappori, 2016). Definition 3 of RES accounts for economies of scale as well as for diseconomies of scale. Diseconomies of scale may arise when the cost of a member of *h*-household is greater than the cost of a single reference person. For instance, for a two-member household consisting of disabled and non-disabled adults, the equivalence scale might be greater than 2. However, *z* for two-member families should not exceed 2 when accounting for economies of scale.

The following modification of RES (3) accounts for economies of scale:

Definition 4. RES for *h*-households respects economies of scale in income if and only if it has the following density function $f_t(z)$

$$f_t(z) = \begin{cases} \frac{f_z(z)}{F_z(h) - F_z(1)}, & \text{for } z \in [1, h] \\ 0, & \text{for } z \notin [1, h] \end{cases}, \tag{5}$$

where *h* is the household size, and $f_z(z)$ and $F_z(z)$ are, respectively, the density function and the distribution function of Z_h (3).

Thus RES respecting economies of scale has the truncated distribution (5) in [0, *h*] interval. The upper truncation point *h,* yielding *income per capita,* specifies the lack of economies of scale. The lower truncation point 1, yielding *income per household*, specifies the maximum economies of scale. It is worth adding that this way of accounting for economies of scale is still valid for other specifications of household attributes, e.g. the numbers of adults and children; the sum equals *h.* According to Definition 3, the positional measures of the truncated distribution may serve as single equivalence scales.

RES is a formalisation of real circumstances which practitioners have inevitably experienced when estimating standard equivalence scales from sizable sample data. Although one can discern the continuous character of income or expenditure distributions and indirect utility functions, it was either ignored or overcome by some arbitrary *ad hoc* assumptions.

For instance, Jackson (1968) calculated equivalence scales for the USA using the food budget shares (FBS) as the household welfare function and presented the following result:

"A *typical* adult living alone requires 36% of the income of a *typical* family of four to attain the same *standard of living* or *welfare level* as the family [our emphases]."

Note that Jackson's expression 'equal welfare level' is misleading, as she used *FBS intervals* in her paper. Moreover, Jackson obtained many expenditures (realisations of X_h and Y, in our terms) within welfare intervals. In other words, she received *expenditure distributions* within welfare intervals. Thus Jackson's equivalence scale of 2.78 is not the definitional *number x_h/y* (2) of expenditure of *two individual households*, but the ratio of *average expenditure* referred to as *the typical* value.

RES is not free from the non-identification problem. According to Definition 3, welfare intervals are necessary for the practical use of RES. It seems reasonable to assume that the subjective ordinal assessments of a household's welfare status ('good', 'rather good', 'average' (neither good nor bad) 'rather bad', and 'bad') are the indicators of underlying welfare intervals. Data on the ordinal measurement of household welfare are available from typical household budget surveys.

To illustrate a possible link between subjective ordinal measurements of welfare and underlying intervals, consider food budget shares (the Engel index) as a welfare measure. Figure 1 shows the kernel estimate of the Engel index density function.
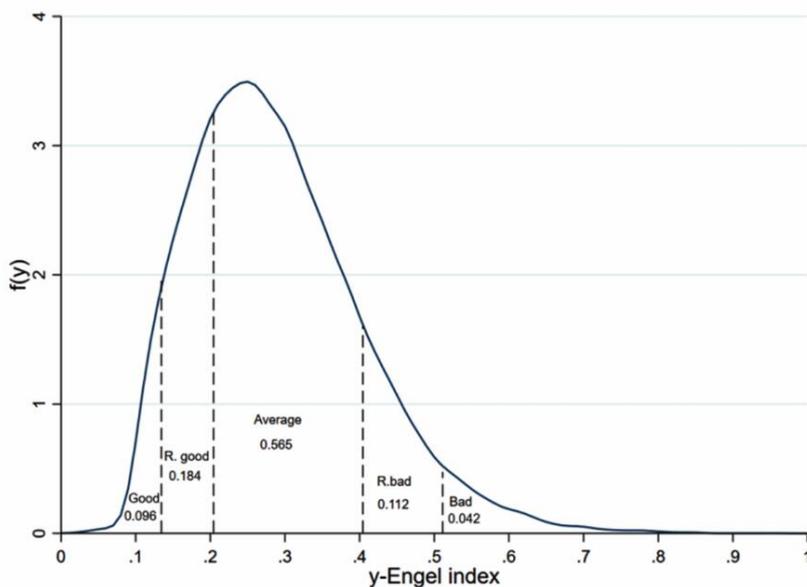

Fig. 1. Welfare intervals corresponding to subjective assessments of welfare status

Source: own elaboration using data from Polish HBS 2015.

By cutting off consecutive probability mass' portions, equal to the observed fractions of subjective welfare measurements, the corresponding welfare intervals were obtained.

The use of subjective assessments of households' welfare has the general restriction that the qualifications: 'good',…, 'bad', have the same meaning to every respondent. Tinbergen (1991) holds that: "this restriction can be accepted since in discussion on the policy resulting from the use of welfare measurements the *same words we also used* either to accept or to reject the policy."

3. RES in the lognormal distribution of income

Let income in the general population of households obey the lognormal distribution with the density function

$$f(x) \; = \; \frac{1}{x\sigma\sqrt{2\pi}}\exp\left\{-\frac{(\ln x-\mu)^2}{2\sigma^2}\right\}, x > 0, \sigma > 0, \tag{6}$$

where μ and σ^2 are the mean and the variance of logarithms of x, respectively (Aitchison and Brown, 1957). The common shorthand $X \sim \Lambda(\mu, \sigma^2)$ was used for a positive random variable X that has the lognormal distribution with parameters μ and σ^2 (Kleiber and Kotz, 2003, p. 107).

The following theorem shows that the lognormal distribution provides a closed form of the density function $f_z(z)$ (4)[5].

Theorem 1. If $X_h \sim \Lambda(\mu_x, \sigma_x^2)$ and $Y \sim \Lambda(\mu_y, \sigma_y^2)$ are independent, then Z_h (3) has the lognormal distribution $\Lambda(\mu_z, \sigma_z^2)$, where $\mu_z = \mu_x - \mu_y$ and $\sigma_z^2 = \sigma_x^2 + \sigma_y^2$.

For proof, see Aitchison and Brown (1957) Theorem 2.3, p. 11.

For the sake of convenience, Z denotes RES for h-household. According to Theorem 1, $Z \sim \Lambda(\mu_z, \sigma_z^2)$, and the density function $f_z(z)$ and the distribution function $F_z(z)$ are, respectively:

$$f_z(z) \; = \; \frac{1}{z\sigma_z\sqrt{2\pi}}\exp\left\{-\frac{(\ln z-\mu_z)^2}{2\sigma_z^2}\right\}, z > 0, \sigma_z > 0, \tag{7}$$

$$F_z(z) \; = \; \begin{cases} 0 \text{ if } z \leq 0 \\ \Phi\left(\frac{\ln z-\mu_z}{\sigma_z}\right), \text{ if } z > 0, \end{cases} \tag{8}$$

---

[5] In mathematics, a *closed-form expression* is a mathematical expression that uses a finite number of standard operations. It may contain constants, variables, and certain 'well-known' operations and functions. The set of operations and functions admitted in a closed-form expression may vary with author and context.

where $\Phi(t)$ is the distribution function of the standard normal distribution. Parameters μ_z and σ_z^2 can be easily estimated from expenditure or income data.

The positional measures of Z are:
the mean

$$M_z = \exp\{\mu_z + \sigma_z^2/2\} \tag{9}$$

the median

$$Me_z = \exp\{\mu_z\} \tag{10}$$

the mode

$$Mo_z = \exp\{\mu_z - \sigma_z^2\}. \tag{11}$$

(Aitchison and Brown, 1957, pp. 8-9). The geometric mean equals the median in the distribution in question.

To derive the truncated distribution (5) of $Z\sim\Lambda(\mu_z, \sigma_z^2)$, where 1 and h are the lower and upper truncation points, these symbols were introduced:

$$\alpha = -\frac{\mu_z}{\sigma_z}, \qquad \beta = \frac{\log h - \mu_z}{\sigma_z}, \qquad \gamma = \Phi(\beta) - \Phi(\alpha),$$

$$\phi(t) = \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}t^2\right\},$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the distribution function and the density function of the standard normal distribution, respectively. The density function $f_t(z)$ and the distribution function $F_t(z)$ of the two-sided truncated lognormal distribution of Z_t are

$$f_t(z) = \begin{cases} \frac{1}{z\sigma_z\sqrt{2\pi}[\Phi(\beta)-\Phi(\alpha)]}\exp\left\{-\frac{(\ln z-\mu_z)^2}{2\sigma_z^2}\right\}, & \text{if } z \in [1,h], \\ 0, & \text{if } z \notin [1,h] \end{cases} \tag{12}$$

$$F_t(z) = \begin{cases} 0, & \text{if } z < 1 \\ \frac{\Phi\left(\frac{\ln z-\mu_z}{\sigma_z}\right)-\Phi(\alpha)}{\Phi(\beta)-\Phi(\alpha)}, & \text{if } 1 \le z < h. \\ 1, & \text{if } z \ge h \end{cases} \tag{13}$$

The rth moment of the truncated lognormal distribution has the form:

$$E[Z_t^r|1 \le z \le h] = E[Z^r]\frac{\Phi(\beta-r\sigma_z)-\Phi(\alpha-r\sigma_z)}{\gamma}, \tag{14}$$

where $E[Z^r] = \exp\{r\mu_z + r^2\sigma_z^2\}$ is the r-th moment of the non-truncated lognormal distribution $Z\sim\Lambda(\mu_z, \sigma_z^2)$ (Jawitz, 2004; Bebu and Mathews,

2009). The second factor in the righthand side of (14) accounts for truncation. It is worth adding that Eq. (14) is valid for all real *r*.

Theorem 2. Let M_z, Me_z, and Mo_z denote the mean, the median, and the mode of the non-truncated distribution of $Z \sim \Lambda(\mu_z, \sigma_z^2)$. Then, the positional measures of the truncated lognormal distribution (10) are:
the mean

$$M_t = c_1 M_z, \tag{15}$$

the median

$$Me_t = c_2 Me_z, \tag{16}$$

the geometric mean

$$G_t = c_3 Me_z, \tag{17}$$

the mode

$$Mo_t = \begin{cases} 1, \text{ if } \mu_z \leq \sigma_z^2 \\ Mo_z, \text{ if } \sigma_z^2 < \mu_z < \ln h + \sigma_z^2 . \\ h, \text{ if } \mu_z \geq \ln h + \sigma_z^2 \end{cases} \tag{18}$$

In Equations (15), (16) and (17), the correcting factors are:

$$c_1 = \frac{\Phi(\sigma_z - \alpha) - \Phi(\sigma_z - \beta)}{\Phi(\beta) - \Phi(\alpha)}, c_2 = \exp\left\{\sigma_z \Phi^{-1}\left(\frac{\Phi(\beta) + \Phi(\alpha)}{2}\right)\right\},$$
$$c_3 = \exp\left\{\frac{\phi(\alpha) - \phi(\beta)}{\Phi(\beta) - \Phi(\alpha)} \sigma_z\right\}.$$

Proof. Formula (15) can be derived from (14) when setting $r = 1$. One can obtain (16) by solving the equation $F_t(Me_t) = 1/2$, where $F_t(\cdot)$ is the distribution function (13) of the truncated lognormal distribution of Z_t. To deriving G_t (17) one may use the fact that $W = \ln Z_t$ has the truncated *normal distribution*, with the density function $g(w) = \phi\left(\frac{w - \mu_z}{\sigma_z}\right) / \sigma_z[\Phi(\beta) - \Phi(\alpha)]$ and the expected value $E_g[W] = \mu_z + \frac{\phi(\alpha) - \phi(\beta)}{\Phi(\beta) - \Phi(\alpha)} \sigma_z$, when the expectation is calculated concerning $g(w)$ (Johnson, Kotz, and Balakrishnan, 1994, p. 10). Then $G_t = \exp\{E_g[W]\}$.

The mode Mo_t (18) is at the maximum of the density function $f_t(z)$ (12). As $f_t(z)$ is a monotonic transformation of $f_z(z)$ (7), the maximum of $f_t(z)$ is the same as that of $f_z(z)$ Three variants of Mo_t (18) are due to whether mode Mo_z of the non-truncated distribution of Z falls into the interval $[1, h]$ or not. QED.

Figures 2a, 2b and 2c illustrate the three variants of Mo_t (18). When $Mo_z \leq 1$, the density function $f_t(z)$ is declining in $[1, 3]$ interval; its maximum equals 1 (Figure 2a). When Mo_z is inside $(1, 3)$ interval, the

density function $f_t(z)$ reaches the maximum at Mo_z (Figure 2b). When $Mo_z \geq 3$, the density function $f_t(z)$ is increasing in [1,3] interval; its maximum equals 3 (Figure 2c).
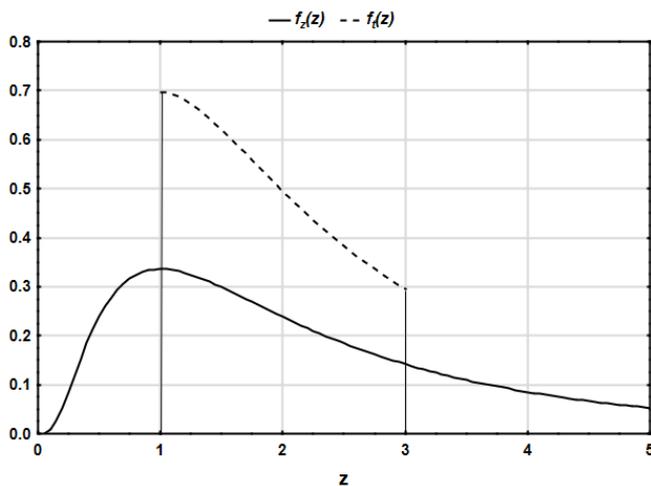

Fig. 2a. Non-truncated (f_z) and truncated (f_t) density functions of RES for three-person households. $Mo_z = 1$
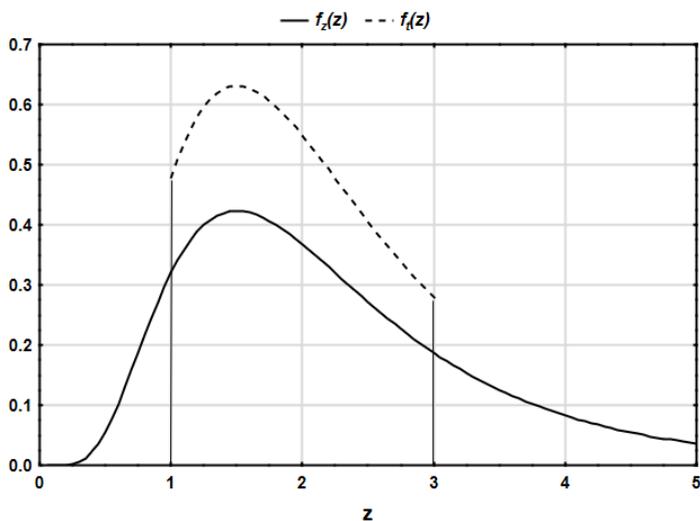
Source: own elaboration.

Fig. 2b. Non-truncated (f_z) and truncated (f_t) density functions of RES for three-person households. $Mo_z = 1.5$
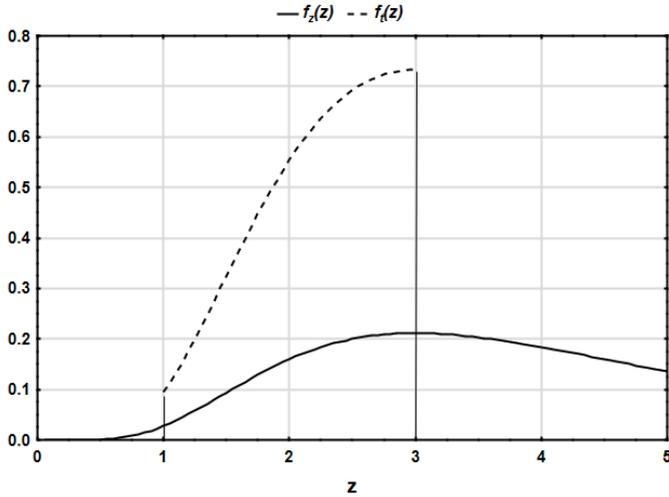
Source: own elaboration.

Fig. 2c. Non-truncated (f_z) and truncated (f_t) density functions of RES for three-person households. $Mo_z = 3$

Source: own elaboration.

4. Normative recommendations for the choice of a single equivalence scale

As mentioned earlier, practitioners need a single equivalences scale. Note that Z (3) would be a *single number* if X_h and $Y's$ distributions were egalitarian (one-point distributions). In reality, however, expenditure distributions are not egalitarian. Nevertheless, one may contemplate a hypothetical egalitarian situation with the concept of *equally distributed equivalent income* (EDEI). EDEI is "[the] level of income which, if distributed equally to all individuals, would generate the same [social] welfare (average utility) as the existing distribution" (Lambert, 2001, p. 95).

Formally, EDEI in the distribution of a positive valued random variable D is the solution, for instance d_ε, to the following equation

$$u(d_\varepsilon) = E[u(D)], \qquad (19)$$

where $u(d)$ is a social planner's utility function. One obtains a single equivalence scale, by construction, when applying (19) directly to the truncated distribution of RES.

Let a social planner assesses the distribution of the random variable D with *the constant inequality aversion function*:

$$u(d) = \begin{cases} \dfrac{d^{1-\varepsilon}}{1-\varepsilon} & \text{for } 0 \leq \varepsilon \neq 1 \\ \log d & \text{for } \varepsilon = 1 \end{cases}, \tag{20}$$

where ε is the measure of inequality aversion (Atkinson, 1970). Then the solution to Equation (19) takes the form:

$$d_\varepsilon = \begin{cases} \{E[D^{1-\varepsilon}]\}^{1/(1-\varepsilon)}, \text{for } 0 \leq \varepsilon \neq 1, \\ \exp\{E[\log D]\}, \text{for } \varepsilon = 1. \end{cases} \tag{21}$$

The distribution of RES may also be the object of ethical judgements, such as distributions of income or expenditure. Note that adjustments of household incomes by distinct equivalence scales yield different income distributions, hence there are various assessments of inequality, poverty, and other distributional issues. Therefore, any ethical evaluation of income distributions embodies an indirect appraisal of underlying equivalence scales.

Let $D = Z_t \sim f_t(z)$, where $f_t(z)$ has the form (12), refer to $z_\varepsilon = d_\varepsilon$ as *the normative equivalence scale*. For calculations of z_ε (21) when $0 \leq \varepsilon \neq 1$, one may apply Equation (14), where $E[Z_t^{1-\varepsilon}]$ is the moment of order $r = 1-\varepsilon$, namely

$$z_\varepsilon = \exp\{\mu_z + (1-\varepsilon)\sigma_z^2/2\}$$

$$\left[\frac{\Phi(\beta-(1-\varepsilon)\sigma_z)-\Phi(\alpha-(1-\varepsilon)\sigma_z)}{\Phi(\beta)-\Phi(\alpha)}\right]^{1/(1-\varepsilon)}. \tag{22}$$

By construction, z_ε (22) is *the normative equivalence scale* preferred by society, which tolerates a given inequality aversion level ε. For $\varepsilon = 1$, z_ε is the geometric mean G_t (17).

By choosing a particular level of inequality aversion ε, one obtains the following specifications of the normative equivalence scales:

$$z_\varepsilon = \begin{cases} M_t, \text{if } \varepsilon = 0 \\ G_t \text{ if } \varepsilon = 1 \\ \text{Proportional to } Mo_t, \text{if } \varepsilon = 3. \end{cases} \tag{23}$$

In (23), M_t, G_t, and Mo_t are the positional measures (15), (17) and (18) of the truncated distribution of RES.

The third variant of (23), with $\varepsilon = 3$, requires explanation. Using (22), the following expression for z_ε is obtained

$$z_\varepsilon = \exp\{\mu_z - \sigma_z^2\}\left[\frac{\Phi(\beta)-\Phi(\alpha)}{\Phi(\beta+2\sigma_z)-\Phi(\alpha+2\sigma_z)}\right]^{1/2}.$$

The first factor in the equation's righthand side is the mode Mo_z (11) of the nontruncated distribution of Z (7) that coincides with the mode

Mo_t (18) of the truncated distribution of Z_t. The second factor plays the role of the coefficient of proportionality. It is easy to see that this factor is less than one. Therefore the normative equivalence scale z_ε for $\varepsilon = 3$ is less than the mode Mo_t.

The result (23) can be interpreted as follows. If society were inequality neutral ($\varepsilon = 0$), it would prefer the equivalence scale based on the mean M_t. The society of a 'moderate' aversion to inequality ($\varepsilon = 1$) would like an equivalence scale equal to the geometric mean G_t. If society tolerated a 'strong' aversion to inequality ($\varepsilon = 3$), the equivalence scale proportional to the mode Mo_t would be preferable. In general, one may use an equivalence scale (22) for any level of inequality aversion $\varepsilon \neq 1$.

5. Empirical results

The study used micro-data on households' monthly expenditures from the Polish Household Budget Survey (HBS) 2015, selecting seven types of households according to various combinations of adults (a) and children (k). The households of single childless persons were the reference type.

Three household welfare categories were specified: 'Good', 'Average' ('neither good, nor bad'), and 'Bad', using the HBS's subjective assessments of households' financial situation. The HBS sample comprised 26,956 households of selected types. The spread ratio test (Thomopoulos, 2017, p. 81) did not reject the null hypothesis that expenditure within each household type follows the lognormal distribution.

Table 1 presents the estimates of single equivalence scales for Poland 2015 within each welfare category. This table contains random equivalence scales based on the mean M_t (15), the geometric mean G_t (17) and the mode Mo_t (18).

Table 1. Random equivalence scales for Poland, 2015

Household's composition	Based on M_t ($\varepsilon = 0$)			Based on G_t ($\varepsilon = 1$)			Proportional to Mo_t ($\varepsilon = 3$)			OECD scales		LIS
	Good	Aver.	Bad	Good	Aver.	Bad	Good	Aver.	Bad	Old	New	
$a + 1k$ ($h = 2$)	1.450	1.451	1.480	1.422	1.423	1.452	1.368	1.369	1.398	1.5	1.3	1.41
$a + 2k$ ($h = 3$)	1.795	1.834	1.812	1.710	1.752	1.732	1.553	1.599	1.583	2.0	1.6	1.73
$a + 3k$ ($h = 4$)	2.225	2.015	2.190	2.083	1.884	2.048	1.825	1.647	1.791	2.5	1.9	2.00
$2a$ ($h = 2$)	1.449	1.457	1.459	1.421	1.429	1.431	1.367	1.375	1.377	1.7	1.5	1.41
$2a + 1k$ ($h = 3$)	1.830	1.863	1.873	1.745	1.780	1.789	1.589	1.625	1.633	2.2	1.8	1.73
$2a + 2k$ ($h = 4$)	2.140	2.162	2.231	1.990	2.019	2.081	1.721	1.761	1.812	2.7	2.1	2.00
$2a + 3k$ ($h = 5$)	2.378	2.397	2.790	2.161	2.198	2.576	1.785	1.848	2.196	3.2	2.4	2.24

Note: h – household size; a – number of adults, k – number of children.

Source: own calculations using data from the Polish Household Budget Survey 2015.

For comparison, Table 1 also contains two parametric equivalence scales commonly used in practice. The first is the OECD scales defined as:

$$s_{ak} = 1 + \theta_1(a - 1) + \theta_2 k, \tag{24}$$

where θ_1, θ_2 are weights assigned to an additional adult and a child, respectively. In the original ('Old') OECD scale, $\theta_1 = 0.7$ and $\theta_2 = 0.5$. The augmented ('New') OECD scale assumes $\theta_1 = 0.5$ and $\theta_2 = 0.3$ (see Hagenaars, de Vos, and Zaidi, 1994; OECD, 1982, 2008).

The second parametric equivalence scale has the form

$$s_h = h^\theta, \tag{25}$$

where h is the household's size, $\theta \in [0,1]$ is a parameter expressing economies of scale (Buhmann et al., 1988). The equivalence scale (25) with $\theta = 0.5$, popularised by the Luxembourg Income Study Group, is known as 'square root' or 'the LIS scale'.

The calculation of the random equivalence scales presented in Table 1 runs as follows. First, we choose a welfare group, e.g. "Good", then select single childless persons' reference households and fit the lognormal distribution, for example, $\Lambda(\mu_a, \sigma_a^2)$ to expenditure distribution. Next, we choose a particular group of households, e.g. households of single persons with one child, and fit the lognormal distribution, say $\Lambda(\mu_{ak}, \sigma_{ak}^2)$ to expenditure distribution. Thus we obtain the distribution of (non-truncated) RES, i.e. $Z_{ak} \sim \Lambda(\mu_{ak} - \mu_a, \sigma_a^2 + \sigma_{ak}^2)$, according to Theorem 1. Finally, the positional measures of the truncated distribution of RES within [1, h] interval are calculated. In the presented example, $h = 2$.

Examining Table 1 shows some interesting features of RES, which depends on welfare. As the welfare qualifications, "Bad", "Average", and "Good" accompany increasing expenditure, one can observe *declining equivalence scales with expenditure*. This observation is consistent with other authors' findings (see e.g. Donaldson and Pendakur, 2003, 2006; Majumder and Chakrabarty, 2008). Thus, the *IB/ESE* assumption does not hold.

Polish households enjoyed significant economies of scale in 2015 since all the estimates were less than the household's size h, reflecting the lack of economies of scale. The random equivalence scales located themselves between OECD Old and OECD New scales. The LIS scale was close to the random equivalence scales based on the geometric mean. Thus the LIS scale seemed appropriate for a society that tolerates a moderate inequality aversion ($\varepsilon = 1$).

The study calculated the cost of having an additional child as random equivalence scales' increments, with a constant number of adults. Table 2 presents the results.

Table 2. The relative cost of additional children in Poland, 2015

Household's composition	Based on M_t ($\varepsilon = 0$)			Based on G_t ($\varepsilon = 1$)			Proportional to Mo_t ($\varepsilon = 3$)		
	Good	Average	Bad	Good	Average	Bad	Good	Average	Bad
$a + 1k$	0.450	0.451	0.480	0.422	0.423	0.452	0.368	0.369	0.398
$a + 2k$	0.345	0.383	0.332	0.288	0.329	0.280	0.185	0.230	0.185
$a + 3k$	0.430	0.181	0.378	0.373	0.132	0.316	0.272	0.048	0.208
$2a + 1k$	0.381	0.406	0.414	0.324	0.351	0.358	0.222	0.250	0.256
$2a + 2k$	0.310	0.299	0.358	0.245	0.239	0.292	0.132	0.136	0.179
$2a + 3k$	0.238	0.235	0.559	0.171	0.179	0.495	0.064	0.087	0.384

Note: a – number of adults, k – number of children

Source: own calculations using data from Table 1.

Table 2 shows that the cost of having an additional child is not constant, as the OECD scale assumes. Generally, the costs diminish with the increase in the number of children. However, in some cases, the cost of having a third child is greater than that of the second, perhaps due to the need to increase the size of dwelling.

6. Conclusions

The general population perspective reveals the stochastic nature of deterministic microeconomic equivalence scales: they become continuous random variables. RES is a formalisation of the real circumstances that practitioners have inevitably experienced when dealing with household populations.

The truncated distribution of RES can account for possible economies of scale in incomes or expenditures. Chiappori (2016) showed that the standard microeconomic equivalence scales fail in this regard.

In applications, the positional measures of the truncated distribution of RES may serve as single equivalence scales. Society's attitude toward inequality may help chose such scales. When income distribution is lognormal (a testable assumption), the calculation of single equivalence scales is easy.

The empirical analysis of SES reveals some interesting facts. Equivalence scales decline with expenditure, thus the *IB/ECE* assumption does not hold, and additionally equivalence scales fall with inequality aversion, which implies increasing economies of scale with inequality aversion.

The presented version of RES has two apparent limitations concerning the measurement of welfare intervals and the lognormal distribution of RES. This paper uses subjective assessments of households' welfare status

as indicators of underlying welfare intervals. One can apply the ordered logit or probit models to estimate the intervals (*cut points*) when using convincing welfare variables (Cameron and Trivedi, 2010). Szulc (2009) proposes several variables as indicators of well-being.

Although the lognormal distribution fits the income data quite well, it might not be true in general. However, for other theoretical models of the income distribution, the density function (4) might have not a compact expression. Pollastri and Zambruno (2010) analysed the ratio of the Dagum distributions numerically, and did not consider a truncated distribution of the ratio.

Acknowledgements

References

Aitchison, J., and Brown, J. A. (1957). *The lognormal distribution with special reference to its uses in economics.* Cambridge: Cambridge University Press.

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, (2), 244-63.

Bebu, I., and Mathew, T. (2009). Confidence intervals for limited moments and truncated moments in normal and lognormal models. *Statistics & Probability Letters*, 79, 375-380.

Bellemare, C., Melenberg, B., and Van Soest, A. (2002). Semi-parametric models for satisfaction with income. *Portuguese Economic Journal*, (1), 181-203.

Biewen, M., and Juhasz, A. (2017). Direct estimation of equivalence scales and more evidence on independence of base. *Oxford Bulletin of Economics and Statistics*, (79), 875-905.

Blackorby, C., and Donaldson, D. (1993). Adult-equivalence scales and the economic implementation of interpersonal comparisons of wellbeing. *Choice and Welfare*, (10), 335-361.

Blundell, R., and Lewbel, A. (1991). The information content of equivalence scales. *Journal of Econometrics*, (50), 49-68.

Bollinger, C. R., Nicoletti, C., and Pudney, S. (2012). Two Can Live as Cheaply as One ... but three's a Crowd. *ISER Discussion Paper* No. 2012-10. Institute for Social and Economic Research, University of Essex.

Buhmann, B., Rainwater, L., Schmaus, G., and Smeeding, T. M. (1988). Equivalence scales, wellbeing, inequality, and poverty: sensitivity estimates across ten countries using the Luxembourg Income Study (LIS) database. *Review of Income and Wealth*, (34), 115-142.

Cameron, A. C., and Trivedi, P. K. (2010). *Microeconometrics Using Stata.* College Station, TX: Stata Press.

Chiappori, P. A. (2016). Equivalence versus indifference scales. *Economic Journal*, (126), 523-545.

Deaton, A., and Muellbauer, J. (1980). *Economics and consumer behavior*. Cambridge: Cambridge University Press.

Donaldson, D., and Pendakur, K. (2003). Equivalent-expenditure functions and expenditure dependent equivalence scales. *Journal of Public Economics*, (88), 175-208.

Donaldson, D., and Pendakur, K. (2006). The identification of fixed costs from consumer behavior. *Journal of Business Economic Statistics*, (24), 255-265.

Ebert, U., and Moyes, P. (2003). Equivalence scales reconsidered. *Econometrica*, (71) 319-343.

Fisz, M. (1963). *Probability theory and mathematical statistics.* New York: Wiley.

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

Jackson, C. A. (1968). Revised equivalence scale for estimating equivalent incomes or budget costs by family Type."*Bulletin* No. 1570-1572 US Department of Labor. Bureau of Labor Statistics.

Jawitz, J. W. (2004). Moments of truncated continuous univariate distributions. *Advances in Water Resources*, (27), 269-281.

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1994). *Continuous Univariate Distributions.* New York: Wiley.

Jones, A., and O'Donnell, O. (1995). Equivalence scales and the costs of disability. *Journal of Public Economics*, (56), 273-289.

Kapteyn, A., and Van Praag, B. (1978). A new approach to the construction of family equivalence scales. *European Economic Review*, (7), 313-335.

Kleiber, C., and Kotz, S. (2003). *Statistical size distributions in economics and actuarial sciences.* Hoboken, New Jersey: Wiley.

Koulovatianos, C., Schröder, C., and Schmidt, U. (2005). On the income dependence of equivalence scales. *Journal of Public Economics*, (89), 967-996.

Lambert, P. J. (2001). *The distribution and redistribution of income.* Manchester-New York: Manchester University Press.

Lewbel, A. (1999). Consumer demand systems and household equivalence scales. In M. H. Pesaran, P. Schmidt, (Eds.). *Handbook of Applied Econometrics*. Volume II: *Microeconomics.* Oxford: Blackwell Publishers Ltd.

Majumder, A., and Chakrabarty, M. (2008). *Estimating equivalence scales through engel curve analysis*. Discussion Paper ERU, no. 05-2008.

Melenberg, B., and van Soest, A. (1995). *Semiparametric estimation of equivalence scales using subjective information*. Tilburg University.

OECD. (1982). *The OECD list of social indicators*. Paris: OECD.

OECD. (2008). *Growing unequal? Income distribution and poverty in OECD countries*. Paris: OECD.

Pollak, R., and Wales, T. (1979). Welfare comparisons and equivalence scales. *American Economic Review*, (69), 216- 221.

Pollak, R., and Wales, T. (1992). *Demand system specification and estimation*. London: Oxford University Press.

Pollastri, A., and Zambruno, G. (2010). Distribution of the ratio of two independent Dagum random variables. *Operation Research and Decisions*, (3-4), 95-102.

Schwarze, J. (2003). Using panel data on income satisfaction to estimate equivalence scale elasticity. *Review of Income and Wealth*, (49), 359-372.

Slesnick, D. (1998). Empirical approaches to the measurement of welfare. *Journal of Economic Literature*, (36), 2108-2165.

Szulc, A. (2009). A matching estimator of household equivalence scales. *Economic Letters,* (103), 81-83.

Thomopoulos, N. T. (2017). *Statistical distributions. Applications and parameter estimates.* Cham: Switzerland: Springer International Publishing.

Tinbergen, J. (1991). On the measurement of welfare. *Journal of Econometrics*, (50), 7-13.

Van Praag, B. M. (1968). *Individual welfare functions and consumer behavior.* Amsterdam: North-Holland Publishing Company.

Van Praag, B. M. (1991). Ordinal and cardinal utility: an integration of the two dimensions of the welfare concept. *Journal of Econometrics*, (50), 69-89.

SZACOWANIE KOSZTÓW UTRZYMANIA DZIECI NA PODSTAWIE LOSOWYCH SKAL EKWIWALENTNOŚCI

Streszczenie: W artykule koszt utrzymania dziecka jest mierzony za pomocą losowej skali ekwiwalentności (RES). Z perspektywy populacji generalnej deterministyczne mikroekonomiczne skale ekwiwalentności ukazują się jako ciągłe zmienne losowe. W niniejszej pracy uzyskano rozkład RES na podstawie lognormalnego rozkładu dochodów. Ucięty rozkład RES umożliwia uwzględnienie korzyści wielkości dochodów lub wydatków. Za pojedyncze skale ekwiwalentności mogą służyć miary położenia w uciętym rozkładzie RES. W wyborze konkretnej skali pomocna może być postawa społeczeństwa wobec nierówności. W pracy oszacowano RES dla Polski w 2015 r. na podstawie mikrodanych z Sondażu Budżetów Gospodarstw Domowych. W owym roku polskie gospodarstwa domowe wykazywały znaczące korzyści wielkości wydatków. Ponadto skale ekwiwalentności malały wraz ze wzrostem wydatków. Zaobserwowano, że koszt utrzymania kolejnego dziecka nie jest stały – malał on wraz ze wzrostem wielkości gospodarstwa domowego.

Słowa kluczowe: skale ekwiwalentności, rozkład lognormalny, korzyści skali, awersja do nierówności.