




# Income-dependent equivalence scales: A fresh look at German micro-data

Jan Marvin Garbuszus<sup>1</sup> · Notburga Ott<sup>1</sup> · Sebastian Pehle<sup>1</sup>  · Martin Werding<sup>1,2</sup>

Received: 22 May 2019 / Accepted: 11 May 2021 / Published online: 27 August 2021  
© The Author(s) 2021

## Abstract

Income inequality and poverty risks receive a lot of attention in public debates and current research. To make income comparable across different types of households, applying the “(modified) OECD scale” – an equivalence scale with fixed weights for each household type – has become a quasi-standard in research. Instead, we derive a base-dependent equivalence scale allowing for scale weights that vary with income, building on micro-data from Germany. Our results suggest that appropriate equivalence scales are much steeper at the lower end of the income distribution than they are for higher income levels. We illustrate our findings by applying them to data on family income differentiated by household types. It turns out that using income-dependent equivalence scales matters for applied research on income inequality, especially if one is concerned with the composition, not just the size of the population at poverty risk.

**Keywords** Equivalence scales · Household income · Household needs · Inequality measurement · Poverty risks

## 1 Introduction

This paper focuses on the problem of making income data comparable across households that differ in size and composition. Equivalence scales are a standard tool that is typically applied in this context (Lewbel and Pendakur 2008). Widely used scales such as the (modified) OECD scale (Hagenaars et al. 1994)<sup>1</sup> consist of scale weights that are fixed (“base-independent”) for each household type across the entire range of income. However, appropriate scale weights are base- or income-dependent, as is suggested by virtually

---

<sup>1</sup>Henceforth, we will call this scale “OECD scale”. It was introduced modifying a scale which was much used earlier and is now called “old OECD scale” or “Oxford scale”.

✉ Sebastian Pehle  
Sebastian.Pehle@ruhr-uni-bochum.de

<sup>1</sup> Ruhr-University Bochum, Faculty of Social Science, Bochum, Germany

<sup>2</sup> CESifo Research Network, Munich, Germany

all empirical estimates allowing for this feature (see below). Resulting equivalence scales mostly decrease with income, so that the scale weight for a couple with one child may be far greater than unity – *i.e.*, the reference weight attached to a single adult – when income is low, but rather close to unity in the range of higher income.

Most of the literature on this issue is based on estimations of “subjective” scales, relating to satisfaction with income or similar perceptions (see, e.g., Koulovatianos et al. 2005; Biewen and Juhasz 2017), but there are also examples for income-dependent scales based on data on income and expenditure (e.g., Donaldson and Pendakur 2004, 2006), like the one suggested here. The equivalence scale derived in this paper is obtained using a two step-approach which combines statistical matching with the estimation of piece-wise linear expenditure systems (Dudel et al. 2017) for four waves of the quinquennial German Sample Survey of Income and Expenditure (EVS) for the years 1998 to 2013. The empirical task we are pursuing is not entirely novel. But we have reasons for adopting a specific – and rather basic – approach, and we think that the implications of income-dependent equivalence scales for inequality measurement have not been sufficiently understood and discussed thus far.

Using equivalence scales with fixed weights when the latter should in fact be income-dependent inevitably leads to biased assessments of poverty and inequality under quite general conditions (Dudel et al. 2021). To illustrate our findings, we therefore apply them to data on family income provided by the German Socio-economic Panel (SOEP) for the period from 1992 to 2015. Comparing our results to those deriving from the OECD scale which is conventionally used for similar purposes (see, e.g., Grabka and Krause 2005; Grabka and Frick 2010; Stichnoth 2016), we show that the choice between equivalence scales really matters. Applying our income-dependent scale increases at-risk-of-poverty (ARP) rates, with consequences that differ considerably by household type. These findings are persistent over the years.

The remainder of this paper is organized as follows. In Section 2, we introduce our empirical approach to estimating base- or income-dependent equivalence scales. Section 3 explains the application to our data and further analytical procedures. Section 4 presents the results and compares them to earlier work. In Section 5, we illustrate the implications and conclude with a discussion of our findings in Section 6.

## 2 Equivalence scales

Empirical research on equivalence scales was intense in the 1970s and 1980s (Kapteyn and Van Praag 1976; Deaton and Muellbauer 1980; Browning 1992; Pollak and Wales 1992; Nelson 1993). More recently, the topic has gained new attention (Blundell et al. 2003; Donaldson and Pendakur 2004; Chiappori 2016), yielding interesting new insights. For applied purposes, however, the simple OECD scale suggested by Hagenaaers et al. (1994) has become a quasi-standard for current research on income inequality and poverty.<sup>2</sup> The OECD scale has no direct empirical basis, but was essentially meant to reconcile diverging empirical findings available at the time of its introduction. It was set up to provide a default

---

<sup>2</sup> The OECD scale attaches a weight of 1.0 to a single adult or to one adult living in a larger household; weights for additional household members are 0.5 (if they are aged 14 and older) or 0.3 (if they are aged less than 14).

scale for analyses lacking more detailed results and, above all, for comparing income across countries.<sup>3</sup>

## 2.1 Expenditure-based scales

The OECD scale belongs to a broader class of equivalence scales that essentially rest on experts' choice and are often meant as normative benchmarks. Equivalence scales can be derived empirically by either comparing subjective perceptions of levels of welfare in households of differing size and composition (see, e.g., Kapteyn and Van Praag 1976) or by analyzing their observed behaviour, most notably their expenses on household consumption (see, e.g., Deaton and Muellbauer 1980). The latter approach is arguably most appropriate for the task of making data on household income comparable with regard to consumption possibilities and material well-being.

Size and composition of households, e.g., the number of persons in a household and their age, can be included in a vector  $\mathbf{z}$  describing household characteristics. Equivalence scales are then meant to measure the difference in income  $y$  that ought to be available to a given household with  $\mathbf{z}_h$  compared to a reference household with  $\mathbf{z}_r$  to provide the two households with the same level of welfare, which is usually represented by a simple household utility function  $u$ .<sup>4</sup>

Household utility is not directly affected by  $y$ , but rather by its use for consumption at given prices  $\mathbf{p}$ . It can therefore be measured by the indirect utility function  $u = V(\mathbf{p}, \mathbf{z}, y)$ , and an expenditure function  $E(\mathbf{p}, \mathbf{z}, u)$  reflecting the resulting demand for goods and services (at minimum costs) can be derived. Estimated equivalence weights  $A_h$  capture the difference between conditional expenditure functions of a household of composition  $\mathbf{z}_h$  and a household of composition  $\mathbf{z}_r$  yielding the same level of  $\bar{u}$ . They are derived as

$$A_h(\mathbf{p}, \mathbf{z}_h, \bar{u}) = \frac{E(\mathbf{p}, \mathbf{z}_h, \bar{u})}{E(\mathbf{p}, \mathbf{z}_r, \bar{u})} = \frac{E_h}{E_r}. \quad (1)$$

Differences between  $E_h$  and  $E_r$  for a given  $\bar{u}$  can be explained by *economies of scale* reflecting advantages of larger households in the consumption of shared goods, such as accommodation or mobility, as well as *economies of scope*, e.g., in the preparation of food, housekeeping or entertainment. Under the assumption that households spend all  $y$  on  $E$ , interpreting savings as future consumption, equivalent income  $y_e$  for a household with  $\mathbf{z}_h$  can be calculated by

$$y_{e,h} = \frac{y_h}{A_h}. \quad (2)$$

Since utility  $u$  is unobservable, a rather strong assumption is necessary to identify equivalence scales from expenditure data, *viz.* that preferences of household members do not change with household characteristics, which means that the functional form of  $u$

<sup>3</sup>Dudel et al. (2020) compare different methods for estimating income-independent equivalence scales. Based on German micro-data, they find that the results are mostly close to the OECD scale, regardless of the precise approach employed.

<sup>4</sup>In a recent paper, Apps and Rees (2020) question this assumption with implications for the use of equivalence scales.

is independent of  $\mathbf{z}$  (Blundell and Lewbel 1991; Lewbel and Pendakur 2008).<sup>5</sup> As this assumption is untestable, it has to be treated as a potential source of measurement error.

Another assumption that is often made to avoid identification problems in estimating equivalence scales is the so called *independence of base* (Lewbel 1989; Lewbel and Pendakur 2008), stating that  $A_h$  does not vary with  $u$  or, in this context, with  $y$  and reduces to  $A_h(\mathbf{p}, \mathbf{z}_h)$ . In other words, equivalence weights for each household type are constant over the whole range of  $y$ . This assumption may be deemed implausible, as additional expenditure for an additional household member can be expected to decrease when  $y$  increases, e.g., because the probability that accommodation needs to be changed is higher if current dwelling size is limited by low income.

## 2.2 Base-dependent equivalence scales

The estimation of equivalence scales based on expenditure data can be traced back to Engel (1857) and Rothbarth (1943) who estimated equivalence weights based on expenditure on single goods.<sup>6</sup> Taking into account household expenditure on more than one good, *Linear Expenditure Systems* (LES) estimate simultaneous equations for different groups of goods (Klein and Rubin 1947; Stone 1954). Further progress was made with the inclusion of savings in *Extended Linear Expenditure Systems* (ELES; Lluch 1973) and of additional socio-demographic parameters in *Functionalized Extended Linear Expenditure Systems* (FELES; Merz 1983).<sup>7</sup>

Common to all these approaches is the assumption of linearity of the marginal effects of increases in income on expenditure. Non-linear expenditure systems such as the *Almost Ideal Demand System* and the *Quadratic Almost Ideal Demand System* (QUAIDS) were suggested by Deaton and Muellbauer (1980) and Banks et al. (1997), respectively. The two latter approaches appear more elegant from a theoretical point of view, but data requirements are much larger, as they require information on (and also some variation in) prices of goods, whereas  $\mathbf{p}$  can be taken to be constant in cross-section surveys for the LES, ELES and FELES. We will use the latter approach for our analyses.

The FELES model developed by Merz (1983) is based on a set of expenditure functions  $a_j$  for commodity groups  $j = 1, \dots, J$ , with

$$a_{jh} = b_{jh} + c_j \left( y_h - \sum_{k=1}^J b_{kh} \right). \quad (3)$$

Here,  $b_{jh} = \sum_{g=1}^G \beta_{jg} z_{gh}$  represents a basic level of consumption of good  $j$  by household  $h$  which is influenced by socio-demographic characteristics  $\mathbf{z}_{gh} = (z_{1h}, \dots, z_{Gh})$ , while  $c_j$  is the marginal budget share of good  $j$  relating to residual income  $y$ , corrected for basic expenditure on other goods. This leads to an equation system, where households may differ

<sup>5</sup>To be sure, this is a strong assumption. However, assuming constant preferences is generally an important ingredient in economic and econometric analyses. Otherwise, virtually anything could be explained by differences in tastes. For alternative approaches, see, e.g., Fleurbaey (2002); a related survey and a practical application are provided in Bargain et al. (2013).

<sup>6</sup> Engel (1857) considered two households to have equal utility  $u$  if they spend the same share of their income  $y$  on food, while Rothbarth (1943) considered the same to be true for two households (with different numbers of children) spending the same absolute amount on so called "adult goods".

<sup>7</sup>Howe (1975) showed that the basic LES is not identified in cross-section analyses without price variation, while – through the inclusion of savings – the ELES is fully identified. FELES retains this property and adds specific assumptions regarding the structure of preferences that further contribute to the identification.

in  $b_j$  but not in  $c_j$ .<sup>8</sup> Equivalence weights can then be derived from the equation system as the ratio of household income  $y_h$  over household income  $y_r$ , with

$$A_h = \frac{y_h}{y_r} = 1 + \frac{\sum_{j=1}^J b_{jh} - \sum_{j=1}^J b_{jr}}{y_r}, \quad (4)$$

*i.e.*, from the difference in basic expenditure  $b_j$  between household type  $h$  and reference households  $r$ .

The FELES approach as such therefore leads to equivalence weights that are independent of income  $y$ . We introduce income-dependency by estimating separate equation systems for households in different income brackets. Households of type  $h$  are subdivided in  $s = 1, \dots, S$  brackets for household income  $y_{hs}$  and compared to meaningful selections of reference households  $r$ . Equivalence weights are assessed for each bracket using

$$A_{hs} = \frac{y_{hs}^*}{y_{rs}^*} = 1 + \frac{\sum_{j=1}^J b_{jhs} - \sum_{j=1}^J b_{jrs}}{y_{rs}^*}. \quad (5)$$

We now write  $y^*$  instead of  $y$ , as we have to transform income data by centering them around the medians for each bracket to arrive at reasonable estimates (see Section 3.3). Income brackets can be defined in various ways. We decided to use quintiles as brackets for households  $h$ , while the selection of reference households requires further consideration (see Section 3.2).

In the following, we therefore run separate estimations not only by survey waves and household types but also by income quintiles. This simple “piece-meal” approach is meant to stay as close as possible with the data and to allow for any kind of variation in equivalence weights by income levels.

### 3 Data and analytical set-up

We estimate equivalence scales using the piece-wise FELES (PFELES) approach based on data from the German Survey of Income and Expenditure (*Einkommens- und Verbrauchsstichprobe*; EVS). Here, we explain how we constructed our data base and how we implemented our model in the empirical work.

#### 3.1 Data pre-processing

The EVS is a quinquennial survey covering about 0.2% of the German population and includes detailed information on income and expenditure as well as some amount of socio-demographic information on households and individuals covered. Variables needed for the estimation are available in survey years 1998, 2003, 2008 and 2013,<sup>9</sup> which we use as a series of cross sections since households can not be tracked across waves, although large overlaps can be assumed due to sampling procedure (Statistisches Bundesamt 2017).

<sup>8</sup>This is the main identifying assumption. A variant is described in Merz (1983) where  $c_j$  is also allowed to vary across households (“FELES 2”). Due to stability issues arising when the number of observations for each household type is small – resulting in large within-type heterogeneity – we decided not to use it in the present paper, although the implementation is straightforward, if data requirements are met.

<sup>9</sup>Data from the latest EVS wave, collected in 2018 and available for researchers since April 2020, are not covered in the empirical work presented here. Given the high stability of our results across survey waves (see Section 4), we would not expect them to convey substantial new insights.

Our analyses concentrate on the following household types: single adults ( $A$ ), childless couples ( $AA$ ), couples with one, two and three children ( $AAC$ ,  $AACC$ ,  $AACCC$ , respectively) and single parents with one child ( $AC$ ), defining the vector  $\mathbf{z}$ . Other household compositions are possible and potentially interesting, but our estimations require a sufficient number of cases, which turns out to be a limiting factor already for  $AACCC$ . From very detailed data on household expenditure, we form broader groups of commodities, encompassing *food and beverages*, *clothing*, *housing*, *media use*, *mobility* and *leisure and education*, and add a further category for *saving, insurance and durables*, assuring separability between groups and avoiding zero expenditure on any of these categories.

In this context, a few aspects deserve further explanation. For instance, to obtain comparable data regarding expenditure on housing for tenants and owners, we use imputed rents for home-owners (as provided by Statistisches Bundesamt 2017), correcting household income accordingly. Similarly, we try to align the way in which insurance contributions or premiums are treated for members of statutory *vs* private health insurance.<sup>10</sup>

A further issue relates to expenditure on consumer durables – e.g., cars, furniture or various household appliances – which can be rather high and accrues only irregularly. Data collection in the EVS refers to a particular quarter of the year, resulting in zero expenditure on these items for the majority of households.<sup>11</sup> We therefore include only “regular” expenditure in the different commodity groups and collect expenditure on durables in a residual group, together with saving, insurance premiums and other precautionary expenses. Expenditure in this residual group varies between 10% and 30% from low-income to high-income households, while the dispersion of expenditure attributed to the other groups is substantially reduced, so that we consider this an acceptable simplification.

We also take measures to reduce heterogeneity among the households considered, focussing on households with adults of working age and with children under age 18, since households with older members can be expected to have different needs, e.g., with respect to health goods and services. We therefore exclude all households with pensioners or with members aged above 65. To avoid misperceptions of atypical constellations, we also exclude couples with more than 10 years of age difference as well as same-sex couples. Furthermore, we exclude households reporting income below the minimum-level of income support (*SGB-II-Regelbedarf*, without cost of accommodation), zero expenditure on food and beverages, or total monthly expenditure above 35.000€, considering these cases as implausible.

Finally, we conduct a multi-variate outlier analysis based on the Cook distance (Cook 1979) of predicted average expenditure,<sup>12</sup> excluding all households whose distance  $d_i^C$  exceeds the 90%-quantile of a central  $F$ -distribution with degrees of freedom  $df1 = L$  and  $df2 = M - L$ , where  $L$  is the number of predictors and  $M$  the number of observations.

### 3.2 Matching households and reference households

To arrive at income-dependent equivalence weights, each of our estimates concentrates not only on a specific household type – paired with another type of reference households – but

<sup>10</sup>In addition, several smaller problems had to be solved for consistently grouping goods across survey waves. The code used for pre-processing our data base and for running our estimates can be downloaded from the website of the journal or will be provided by us upon request.

<sup>11</sup>This creates a potential problem for the estimations (Keen 1986).

<sup>12</sup>Effectively, the Cook distance measures how much the average would change, if a single observation were removed from the sample.

also on sub-groups defined by income quintiles (see Section 2.2). Therefore, we have to select appropriate reference households for each of these sub-groups and, at the same time, need to make sure that this selection does not pre-determine our results. For instance, we cannot simply pick parallel income quintiles among reference households, since we do not know beforehand – *i.e.*, before assessing equivalence weights – which ranges of income are indeed comparable across the two types.

We thus use statistical matching at this stage as a kind of formal pre-processing for the following regressions (Dudel et al. 2017). This step is also meant to make households and reference households more homogenous in terms of their preferences and the extent to which they may benefit from economies of scale in consumption. For this purpose, we calculate a distance matrix using the Gower distance (Gower 1971), based on socio-demographic characteristics such as *education, employment status, age, home ownership, balance of assets and liabilities* and the *share of expenditures on food and on clothing*.<sup>13</sup> The Gower distance  $d_{ij}^G$  between household  $i$  and household  $j$  is the average of variable-specific distances  $d_{ijm}$ , where subscript  $m$  denotes the  $n = 1, \dots, N$  variables in

$$d_{ijn}^G = \frac{1}{N} \sum_{n=1}^N d_{ijn}, \quad (6)$$

with  $d_{ijn} = \frac{|n_i - n_j|}{\max\{|n_i - n_j|, \forall i, j\}}$ , if  $n$  is a quantitative variable, and  $d_{ijn} = G(n_i, n_j)$ , if  $n$  is a dummy variable (being equal to 1 if  $n_i = n_j$  and 0 otherwise). As a matching algorithm we use 1:1 optimal matching (Hansen and Klopfer 2006) as it guarantees the groups of households considered and reference households are of equal size.

Ideally, one would like to match households of all other types to just one type of reference households – for instance, AA –, to obtain a direct link between all the estimates for equivalence weights. However, this does not work empirically. Matching between “neighbouring” types, such as AACC and AAC, leads to better results than when matching AACC to AA, as the former are more similar in terms of patterns of labour force participation, expenditure and other matching variables. Also, the direction of matching should be chosen so that the household type with the smaller number of observations is matched to the one with the larger number to alleviate finding matches for all households in the first group. Table 1 gives an overview regarding pairs of households and reference households as well as the direction of matching (“*dom*”) actually chosen. It also indicates income thresholds that were used for a pre-selection of candidate matches.

It was already explained that type- $h$  households are sub-divided into (within-type) income quintiles  $Q_s = Q_1, \dots, Q_5$ , with lower and upper limits  $Q_{ll}$  and  $Q_{ul}$ , respectively. Reference households  $r$  are then selected for the matching process based on the plausible idea that income thresholds for the smaller household type should be lower, but not higher, than for the larger household. We thus start from  $Q_{ll}$  and  $Q_{ul}$  and adjust one of these thresholds for type- $r$  households incrementally in 500€ steps, until we obtain a completely matched sample (see, again, Table 1 for final income thresholds).

<sup>13</sup>Note that expenditure shares for food and clothing could be considered as simple welfare indicators (following Engel 1857; Rothbarth 1943, see footnote 6), while the EVS data do not include any more direct measures.

**Table 1** Direction of matching and relevant income thresholds

$r$	$dom$	$h$	$Q_{ll}$	$Q_{ul}$
A	←	AA	$Q_s - 10,000\text{€}$	$Q_s$
AA	←	AAC	$Q_s - 2,000\text{€}$	$Q_s$
AAC	→	AACC	$Q_s$	$Q_s + 2,500\text{€}$
AAC	←	AACCC	$Q_s - 5,000\text{€}$	$Q_s$
A	←	AC	$Q_s - 2,000\text{€}$	$Q_s$

Household types are denoted by the number of adults (“A”) and children (“C”) living together.  $r$  = reference households;  $dom$  = direction of matching;  $h$  = household type considered;  $Q_{ll}$ ,  $Q_{ul}$  = lower and upper limit of income quintiles;  $Q_s$  = quintiles ( $s = 1, \dots, 5$ )

### 3.3 Regressions and assessment of equivalence weights

We then apply the PFELES approach described in Section 2 and estimate the equation system for differing pairs of households  $h$  and  $r$ .<sup>14</sup> Here, the differentiation by income quintiles requires another transformation. Since basic expenditure  $b_{jhs}$  and  $b_{jrs}$  are effectively vertical intercepts of linear equations (with identical slope  $c_{js}$ ) – so-called “Engel curves” –, we have to center  $y$  from each income bracket around the respective median  $v$ , obtaining transformed values  $y^* = y_h - v$  to arrive at reasonable estimates.<sup>15</sup> As suggested by Merz (1983), each equation system is simultaneously estimated using seemingly unrelated regressions (SUR; Zellner 1962).

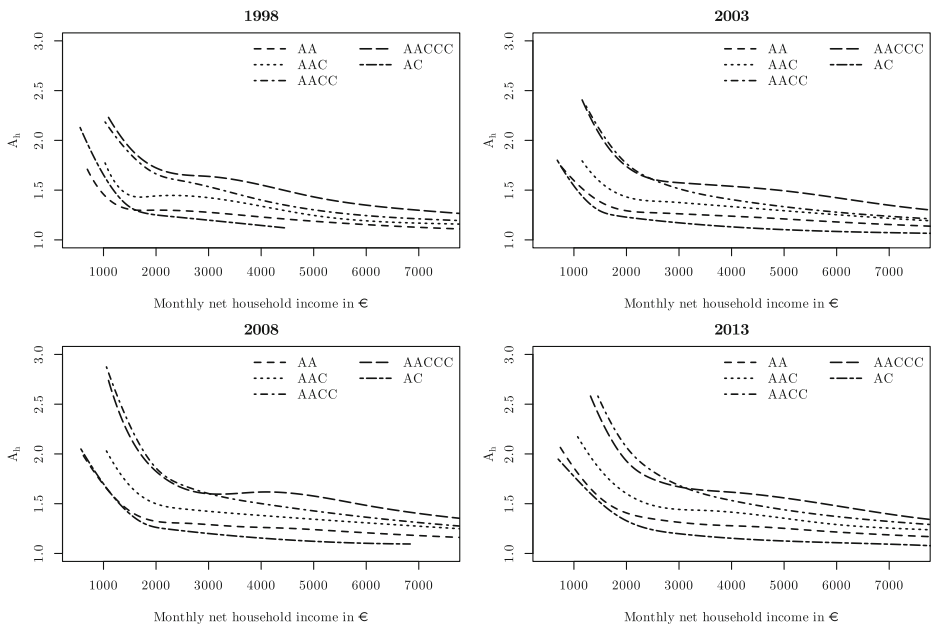
To alleviate interpretation and application, the resulting quintile-wise scale weights are chained (via multiplication) across household types, making sure that type A is the final reference household with an equivalence weight of 1, while all resulting scale weights point to income levels which, compared to a single adult, lead to the same levels of material well-being. Last but not least, base-dependent scale weights for all other household types are smoothed over the entire range of possible incomes  $y$  via natural splines.

When preparing our final empirical approach, we explored a number of alternatives regarding several analytical steps, which proved the robustness of our results. For instance, instead of income quintiles we also looked at terciles and deciles. As can be expected, the results had a less pronounced income-related structure on the one hand and ran into more severe problems with the number of observations on the other. To the extent that they could still be considered reliable, however, they showed a very similar picture as the one provided here. Furthermore, our results do not appear to be rather sensitive to the precise rules applied to eliminating outliers as well as to the matching algorithm used and the distance measure involved. The application of statistical matching, the choice of matching variables, and the

<sup>14</sup>Note that in the empirical application, the identifying assumption of the FELES model, *viz.* that marginal expenditure on each category of consumption goods,  $c_j$ , is identical (see Section 2.2), is taken to apply to each pair of household types, not to all household types considered at the same time. This is a conventional relaxation of the underlying theoretical restrictions for the estimation of base-independent scales. To arrive at income-dependent scales, we also allow  $c_j$  to vary with income for any given pair of household types.

<sup>15</sup>Otherwise, estimates for  $b_{jhs}$  and  $b_{jrs}$  would also depend on how far average income in each bracket is away from the origin of the income axis.





**Fig. 1** PFELES equivalence scales for monthly net household income. Sources: EVS (1998, 2003, 2008, 2013), own calculations. Notes: Household types are denoted by the number of adults (“A”) and children (“C”) living together. Each panel shows income-dependent equivalence weights,  $A_h$ , by household types (compared to a single adult) for one of the quinquennial survey waves

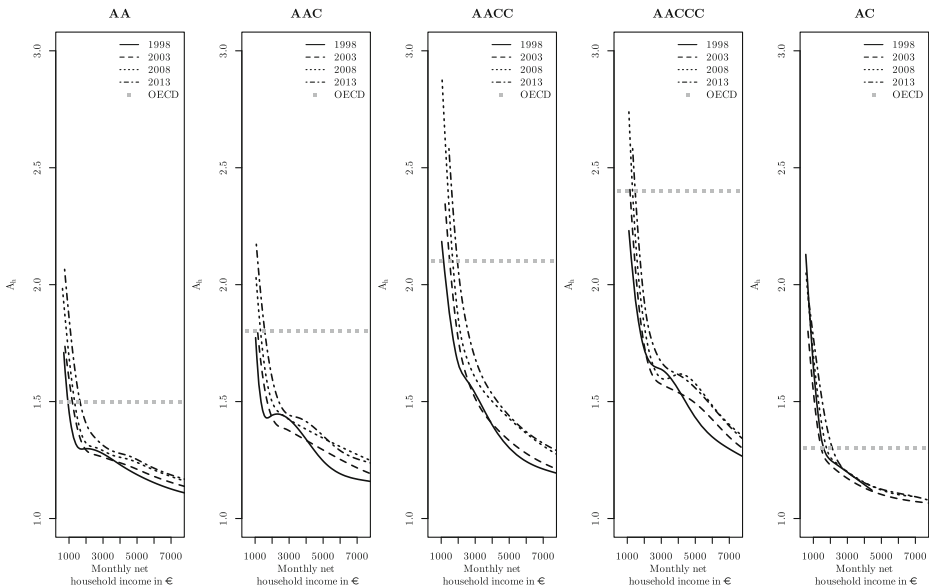
centering of income data to focus on local differences in basic expenditure in each income bracket appear to be important aspects of our approach.

## 4 Results

The equivalence scales resulting from the procedures described in Section 3 are shown in Fig. 1. As can be seen, the shape and structure of these scales can be considered as highly plausible, mostly leading to (1) monotonically decreasing equivalence weights as income increases for each household type and (2) increasing equivalence weights as household size increases.

Against this background, anomalies occur (i) for AAC in survey year 1998 and AACCC in survey year 2008, where scale weights locally show slight increases for intermediate levels of income, (ii) for AC in 1998, where estimated weights lie above those for AA for very low levels of income (while they are rather close to those for AA in the other survey years), and (iii) for AACCC, where scale weights are close to, and even below, those for AACC for low levels of income in three of our four survey years.

Observation (ii) does not necessarily point to an inconsistency of our estimates. Needs of a child may differ so much from those of an adult that they require additional expenditure which is almost as high or even higher than that for a second adult, if income is very low. While the situation for households with larger numbers of children may be similar, we take our results for household type AACCC on low levels of income to be implausible, certainly with respect to survey year 2013. We mainly attribute anomaly (iii) to an



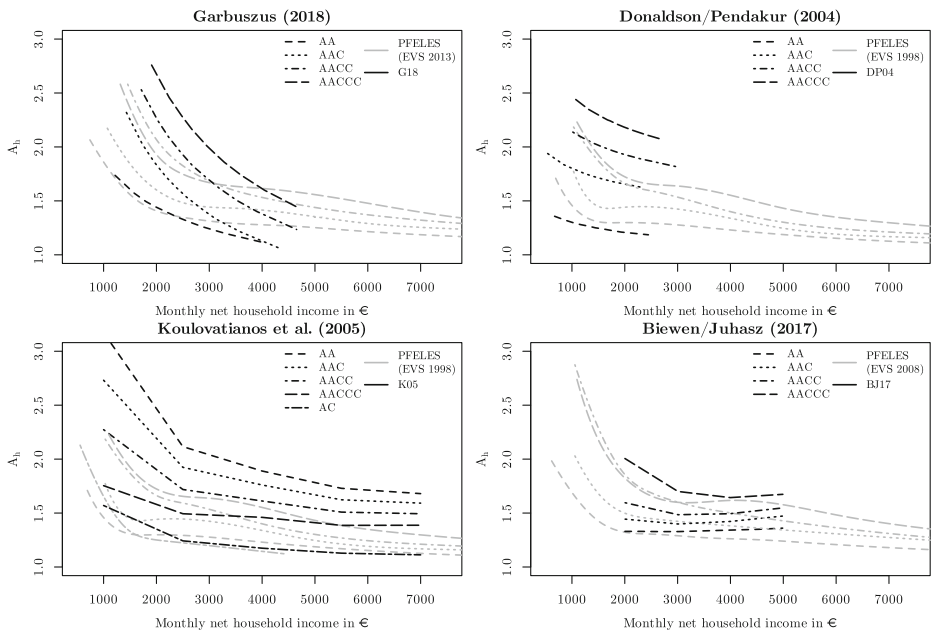
**Fig. 2** PFELES equivalence scales by household type. Sources: EVS (1998, 2003, 2008, 2013), own calculations. Notes: Household types are denoted by the number of adults (“A”) and children (“C”) living together. Each panel shows income-dependent equivalence weights,  $A_h$ , by survey waves for a given household type (compared to a single adult). Grey horizontal lines represent OECD scale weights (based on constant weights of 0.3 for a child regardless of age)

insufficient number of observations. The number of couples with three (or more) children is decreasing substantially from survey year to survey year.<sup>16</sup> In addition, we also observe problems at the matching-stage for this household type, as patterns of labour force participation in type-*AACCC* households tend to differ from those of couples with smaller numbers of children.<sup>17</sup>

Otherwise, the estimated scale weights exhibit a very clear shape. The weights are rather high – well above those of the OECD scale (see Fig. 2) – at the bottom end of the income distribution, but they are decreasing rapidly as income increases and become rather flat from a medium range of income onwards. In other words, economies of scale and scope that are behind the form of these curves appear to be weak at low levels of income, mostly driven by higher expenditure on food and housing, while they become substantial when household income exceeds a certain threshold. As a consequence, scale weights differ a lot for low-income households, while they tend to converge to rather low figures for high-income households.

<sup>16</sup>While there are more than 200 observations per income quintile for *AACCC* in 1998, there are less than 80 in 2013. We therefore decided to pool observations from all four waves for this household type; the variation across waves shown in Fig. 1 is thus the result of linking identical results for *AACCC* to scales weights for *AAC* and *AA* that vary slightly over time.

<sup>17</sup>Again, this difference becomes larger over time, mainly because labour-force participation of female adults does not change much in *AACCC*-households. It may be worth mentioning that matching *AACCC* to *AACC* (instead of *AAC*)-households produces a quality of matches and estimates of scale weights that look even less plausible.



**Fig. 3** PFELES scales compared with other base-dependent scales. Sources: Garbuszus (2018, figure 9.1), Donaldson and Pendakur (2004, figure 2), Koulovatianos et al. (2005, figure 1) and Biewen and Juhasz (2017, figure 7). Notes: Household types are denoted by the number of adults (“A”) and children (“C”) living together. Each panel shows income-dependent equivalence weights,  $A_h$ , by household types (compared to a single adult) deriving from our estimates (grey lines) and from estimates run by other authors (black lines)

Interestingly, there is little variation in scale weights across the different survey years (see, again, Fig. 2). Scales mostly appear to show slight horizontal shifts along the y-axis, along with nominal and real income growth (where the latter has been relatively moderate in the relevant period of time). Given our separate estimations, we take this as evidence for the robustness of our approach, while larger variation could point to major changes in household needs or household income, lack of reliability of our specification, problems with the estimations, or any combination of these.

Our piece-meal approach with numerous separate estimates thus confirms that the decline of scale weights with income is not an artifact of assumptions regarding any pre-defined overall shape of expenditure functions. Among other things, this justifies the use of analytical approaches that are more comprehensive and more elegant – and could also improve on weaknesses of our results. In fact, Garbuszus (2018) provides an example, deriving income-dependent scale weights from pooled data of all EVS waves from 1998 to 2008, which slackens data limitations, and employing the non-linear QUADS approach suggested by Donaldson and Pendakur (2004). The results look much like those presented here, but are in line with reasonable expectations also for AACC households across the entire range of income (and cover households of type ACC as well).

In Fig. 3, several income-dependent equivalence scales are compared to our results.<sup>18</sup> While Donaldson and Pendakur (2004) apply the same approach as Garbuszus (2018) to Canadian data, Koulovatianos et al. (2005) and Biewen and Juhasz (2017) derive subjective

<sup>18</sup>The respective scales were extracted from the original papers via WebPlotDigitizer Version 4.3.

scales from data on income satisfaction for the German population. The variation in scale weights for low-income households turns out to be similar across all these papers. At the same time, the slope – and therefore the degree of income dependence – of our scales is steeper than in Donaldson and Pendakur (2004), but a bit less steep than in Garbuszus (2018). This is remarkable, as the two contributions are using the same methodology, so that this may point to cross-country differences with respect to expenditure behaviour as well as socio-economic characteristics of different household types.

Subjective scales are conceptually different from expenditure-based scales. Nevertheless, apart from a smaller variation by household types across the entire range of income, our results are rather similar to those obtained by Koulovatianos et al. (2005). Also, our PFELES weights roughly match with those provided by Biewen and Juhasz (2017) – with the exception of *AACC*-households –, even though they find that scale weights are not monotonically decreasing for several household types.

All in all, there may thus be some way to go to arrive at a suitable methodology and reliable results with respect to income-dependent equivalence scales, but some consensus regarding the shape and structure of such scales has already been established.

## 5 Illustration

To illustrate the implications of income-dependent scales for measuring income inequality and specifically poverty risks, we finally apply the PFELES scales presented in Section 4 to income data from the longitudinal German Socio-economic Panel (SOEP; Goebel et al. 2019) for the years since German reunification.<sup>19</sup>

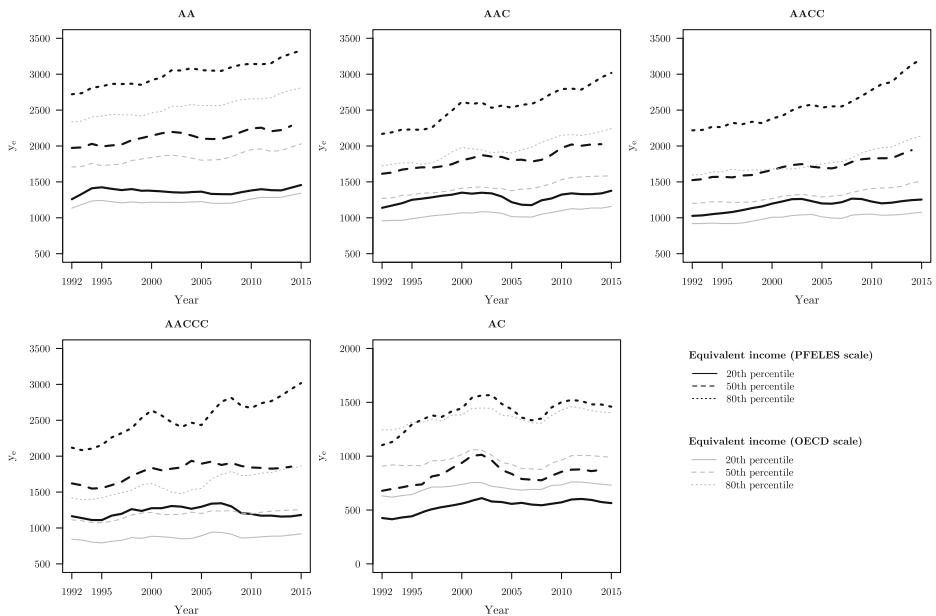
Mirroring restrictions of the EVS sample (see Section 3), we concentrate on the six household types explicitly covered in our estimates and on households with adults of working age and children under age 18. €-specific PFELES scale weights estimated for different waves of the EVS are interpolated via linear splines for intermediate years, while we extrapolate estimates for 1998 and 2013 to the years from 1992 to 1997 and to 2014 and 2015. For SOEP income data outside the range observed in the EVS, we apply the respective minimum or maximum scale weights. Being concerned about the plausibility of our results for household type *AACCC*, we use type-*AACC* scale weights for this type, if the former fall below the latter for low levels of income.<sup>20</sup> Throughout, equivalent income is adjusted to year-2010 prices to correct for inflation.

Figure 4 shows the development of resulting equivalence incomes at three prominent positions in the income distribution of each household type, *viz.*, the 20th, 50th and 80th percentile.<sup>21</sup> For comparison, the figure also displays results for equivalence incomes that are obtained using the OECD scale. It is interesting to see that – with exceptions for type-*AC* households – lines representing equivalent-income percentiles deriving from our scale weights are all above those deriving from the OECD scale, with differences that are smallest at the 20th percentile and increase with income. By contrast, lines for the 20th and 50th percentile of type *AC* are clearly below those based on the OECD scale.

<sup>19</sup>For further illustrative results relating to the income situation of families since the 1990s, see Garbuszus et al. (2018b, ch. 7) or Garbuszus et al. (2018a).

<sup>20</sup>This can be taken to be a lower bound for plausible weights. Still, results for this household type should be considered with a grain of salt.

<sup>21</sup>All graphs included in Figs. 4 through 5 have been smoothed by central moving averages over adjacent years.



**Fig. 4** Median, upper and lower quintile of equivalent monthly net household income. Sources: SOEP v32.1, own calculations. Notes: Household types are denoted by the number of adults (“A”) and children (“C”) living together. Each panel shows the evolution of different percentiles of equivalent income,  $y_e$ , for a given household type based on equivalence scales deriving from our estimates (black lines) and on the (modified) OECD scale (grey lines)

The mechanics behind these observations that also matter for our further illustrations are as follows. Income-dependent PFELES weights for very low levels of income are all above the corresponding fixed OECD weights. They decrease as income increases and fall below the OECD weights mostly in the first or second quintile (see, again, Fig. 2). Only for single parents with one child, an intersection is reached at relatively high levels of income, given the income distribution of this household type. Now, if OECD weights are “too high” for the majority of households, *i.e.*, those with medium and higher income, applying this scale generally compresses the distribution of equivalent income. Based on the PFELES weights, the distribution becomes wider and median equivalent income rises.<sup>22</sup> At the same time, the OECD scale tends to overestimate equivalent income of households on low income compared to the PFELES weights, with effects that differ across household types. Taken together, this contributes to a change in the dimension of income inequality as well as in the size and the composition of the population close to, or at, risk of poverty.

Implications for widely used inequality indices are shown in Table 2. Here, we are looking at simple quintile-ratios (relating to the 80th and 20th percentile of distributions of equivalent income) and several variants of generalized Gini (S-Gini) and generalized

<sup>22</sup>Combining all household types considered here, the median of monthly equivalent income for 2014 is 1,667€ based on the OECD scale and 2,030€ based on the PFELES weights, reflecting an increase by roughly 20%. Note that this has little to do with the fact that we are considering only a sub-set of all households covered in the SOEP data (since we have no scale weights for other types of households). Median equivalent income for the entire SOEP population in 2014 is assessed to be 1,712€ by Goebel et al. (2016) using the OECD scale.

**Table 2** Inequality indices for equivalent monthly net household income

Year	Household	80th/20th pc.		S-Gini ( $\nu = 2$ )		S-Gini ( $\nu = 3$ )		S-Gini ( $\nu = 4$ )		GE ( $\alpha = -1$ )		GE ( $\alpha = 0$ )		GE ( $\alpha = 1$ )		GE ( $\alpha = 2$ )	
		PFELES	OECD	PFELES	OECD	PFELES	OECD	PFELES	OECD	PFELES	OECD	PFELES	OECD	PFELES	OECD	PFELES	OECD
1995	AA	1.98	1.94	0.27	0.25	0.37	0.34	0.43	0.40	0.15	0.12	0.13	0.11	0.15	0.12	0.22	0.17
2005	AA	2.24	2.09	0.29	0.27	0.41	0.37	0.47	0.43	0.21	0.15	0.16	0.12	0.15	0.12	0.19	0.15
2015	AA	2.22	2.05	0.30	0.26	0.41	0.37	0.48	0.43	0.21	0.14	0.16	0.12	0.16	0.12	0.20	0.15
1995	AAC	1.85	1.86	0.26	0.24	0.36	0.33	0.41	0.38	0.14	0.11	0.12	0.10	0.13	0.10	0.16	0.12
2005	AAC	2.16	1.89	0.27	0.24	0.38	0.33	0.44	0.39	0.17	0.11	0.13	0.10	0.13	0.10	0.15	0.11
2015	AAC	2.11	1.91	0.29	0.25	0.40	0.34	0.47	0.40	0.20	0.12	0.15	0.10	0.15	0.10	0.18	0.12
1995	AACC	2.23	1.83	0.28	0.22	0.38	0.30	0.43	0.35	0.14	0.08	0.13	0.08	0.14	0.08	0.19	0.10
2005	AACC	2.12	1.70	0.29	0.22	0.39	0.30	0.45	0.35	0.18	0.09	0.15	0.09	0.17	0.10	0.24	0.13
2015	AACC	2.52	1.98	0.31	0.24	0.42	0.33	0.48	0.38	0.21	0.10	0.16	0.09	0.17	0.10	0.22	0.13
1995	AACCC	1.99	1.86	0.28	0.22	0.37	0.30	0.42	0.34	0.14	0.08	0.13	0.08	0.14	0.08	0.18	0.10
2005	AACCC	1.96	1.75	0.24	0.19	0.34	0.28	0.40	0.33	0.12	0.07	0.10	0.06	0.10	0.07	0.12	0.07
2015	AACCC	2.56	1.96	0.32	0.25	0.43	0.34	0.49	0.39	0.21	0.11	0.17	0.10	0.17	0.10	0.21	0.12
1995	AC	2.89	2.00	0.33	0.24	0.45	0.33	0.51	0.38	0.20	0.10	0.18	0.09	0.18	0.09	0.21	0.10
2005	AC	2.38	1.90	0.29	0.22	0.41	0.31	0.47	0.36	0.17	0.09	0.15	0.08	0.14	0.09	0.17	0.10
2015	AC	2.69	1.94	0.31	0.22	0.42	0.31	0.48	0.36	0.17	0.09	0.15	0.08	0.15	0.08	0.18	0.09
1995	ALL	2.15	2.02	0.29	0.26	0.39	0.36	0.46	0.41	0.17	0.13	0.15	0.11	0.15	0.12	0.21	0.16
2005	ALL	2.31	2.12	0.30	0.27	0.41	0.37	0.48	0.43	0.21	0.14	0.16	0.12	0.16	0.13	0.21	0.16
2015	ALL	2.61	2.28	0.33	0.28	0.45	0.39	0.52	0.45	0.24	0.16	0.19	0.14	0.19	0.14	0.25	0.20

Sources: SOEP v32.1, own calculations

Household types are denoted by the number of adults (“A”) and children (“C”) living together. Indications in the table refer to income ratios between households at the 80th and 20th percentile (80th/20th pc.) of the relevant income distribution; to generalized single-parameter Gini coefficients (S-Gini,  $\nu$  being a parameter measuring inequality aversion); and to generalized entropy measures (GE,  $\alpha$  being a parameter measuring sensitivity to different segments of the income distribution). Results for ALL households encompass single adults and all other household types listed in the table

entropy (GE) indices (thoroughly described in Cowell 2011; Donaldson and Weymark 1980). The table displays results for selected years and for single household types as well as for all household types for which we have been able to estimate income-dependent equivalence weights. With just one exception, inequality increases through the more differentiated PFELES weighting compared to an application of the OECD scale, with differences that vary substantially across household types.<sup>23</sup> Results for *all* households covered here consistently increase by between 5% and 50%. Differences tend to rise over time. Furthermore, GE-measures tend to be more sensitive to the equivalence scales applied than quintile-ratios or S-Gini indices.

A key figure in current research on poverty is the at-risk-of-poverty (ARP) rate. It is defined as the share individuals that have less than 60% of median equivalent net income at their disposal. This concept of relative poverty – dating back to Fuchs (1965) who had suggested a 50%-threshold<sup>24</sup> – can be debated, *inter alia*, due to the arbitrary nature of the thresholds applied. To demonstrate the effects of different equivalence scales and to compare results relating to different thresholds for relative income, Fig. 5 shows the development of ARP rates – or, strictly speaking, of ARP rates (60%) and poverty rates (50%) – for equivalent income based on the PFELES scale and the OECD scale.

Figure 5 indicates that the impact of applying different thresholds for (at risk of) poverty rates is sizable. However, differences relating to the equivalence scales used are no less important.<sup>25</sup> With an exception for type-*AACCC* households, applying the PFELES scale consistently leads to higher ARP rates compared to the OECD scale, with differences that vary substantially by household type.<sup>26</sup> In addition, ARP rates for each household type exhibit differing time trends and, occasionally, differences relating to the use of different scales are far from parallel even for a given household type. Probably the most striking result relates to the situation of type-*AC* households. Single parents with one child appear to be far worse off in terms of poverty (risks) than has ever been shown using the OECD scale (which gives this household type a particularly low weight). Using this scale, the ARP rate is fluctuating around 40% throughout the observation period, while it is around no less than 65% when applying the PFELES scale.

## 6 Discussion

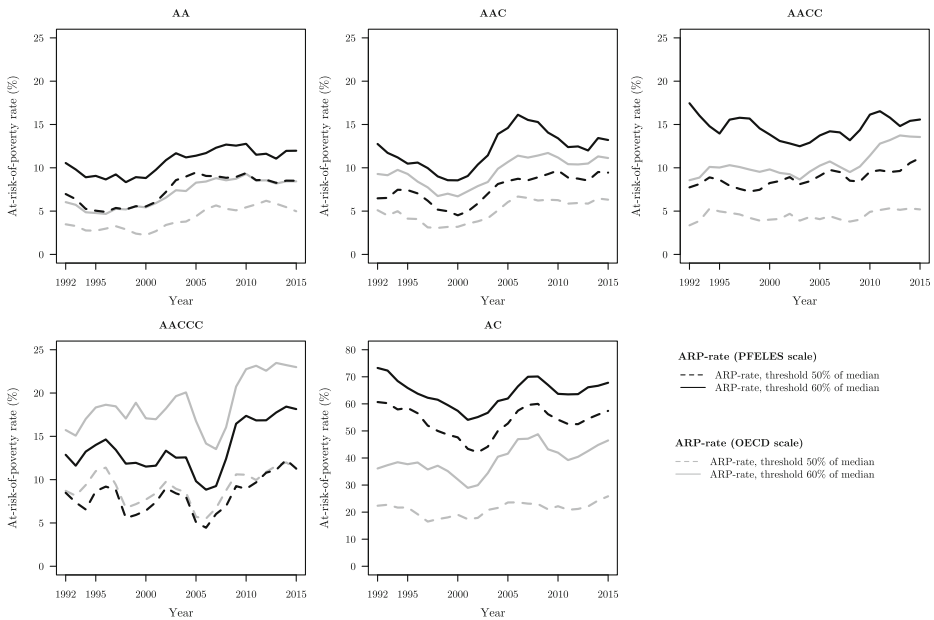
The results presented here support the idea that, if one allows for this kind of variation in the empirical assessment, equivalence scales based on household expenditure vary substantially with household income. They also demonstrate that, neglecting this property, the widely used OECD scale applies scale weights which are particularly inadequate for ranges of

<sup>23</sup>The exception is given by results relating to the quintile-ratio for household type *AAC* in 1995. With the GE-measures for household type *AC*, results based on the PFELES scale can become twice as high as those deriving from the OECD scale.

<sup>24</sup>Poverty research in the US mostly sticks to the 50%-threshold. In Europe, the switch to a 60%-threshold reflecting “poverty risks”, but not necessarily “poverty”, was meant to indicate a more careful use of notions of relative poverty. In practice, however, poverty risks and poverty are now often confused.

<sup>25</sup>Across all our household types, thresholds at 60% of median income for the years of 1995, 2005 and 2015 are 800€, 865€ and 954€ for OECD-weighted income, whereas they are 982€, 1,048€ and 1,168€ for PFELES-weighted income.

<sup>26</sup>Remember that for couples with three children on low income, we obtain implausibly low scale weights. Also, by its simple, additive structure, the OECD scale attaches relatively high weights to this household type. So weaknesses of the two scales probably coincide here.



**Fig. 5** At-risk-of-poverty (ARP) rates by household types. Sources: SOEP v32.1, own calculations. Notes: Household types are denoted by the number of adults (“A”) and children (“C”) living together. Each panel shows the evolution of ARP rates for a given household type based on different thresholds for poverty or poverty risks (50% and 60% of median equivalent income) and on equivalence scales deriving from our estimates (black lines) and on the (modified) OECD scale (grey lines)

income where poverty thresholds are usually located. This leads to distorted results regarding overall measures of income inequality and specifically regarding the level of poverty and/or the structure of households at poverty risk.

We conclude that the impact of income-dependent equivalence scales is important, even though we do not claim to offer a fully-fledged solution regarding approaches for, and final results of, assessments of scales of this type. But then, what are the practical implications of our findings? Do we recommend to replace the OECD scale with income-dependent equivalence scales in all future work on inequality and poverty? The answer is both yes and no. The answer is (partly) “yes” because the OECD scale rests on over-simplifications that render it particularly ill-suited for measuring poverty lines and assessing poverty rates – an area where it has become an international quasi-standard in recent years.<sup>27</sup> But the answer must be (partly) “no” because it is impossible to replace the OECD scale, or alternatives which are similarly simple, in all of their current applications.

The OECD scale was suggested mainly as a compromise meant to be used for cross-country comparisons in cases where more detailed and generally agreed results were lacking (see Hagenaars et al. 1994, p. 194). In this role, simple defaults like the OECD scale may be indispensable. Even for work focusing on just one country, it is all but easy to replace

<sup>27</sup>Interestingly, in a methodological note that is still valid the OECD confronts the “modified” OECD scale with the “old” one and with the “square-root” scale and concludes: “In general, there is no accepted method for determining equivalence scales, and no equivalence scale is recommended by the OECD for general use” (OECD 2021).



it, since simplicity has its merits also at this level. An advantage of the OECD scale is that it is represented by essentially three figures relating to three types of household members (see footnote 2), while income-dependent scales like the one used here need to be spelt out in long arrays of scale weights relating to different household types and, in addition, to different levels of income – in some higher or lower resolution.

A related advantage of the OECD scale is that, due to its “Lego brick” logic, it can be applied to all types of households included in a given population, whether this may be appropriate or not. Income-dependent scales have to be determined specifically for each household type, which almost inevitably runs into problems with data limitations for a non-negligible percentage of the population.

As long as there is mainly an interest in aggregate figures and trends – e.g., ARP rates among the entire population and their development over time – the OECD scale may therefore have to do. Aggregate rates of this kind may be underestimated considerably using the OECD scale compared to our income-dependent scale, but the time trends are practically unaffected over more than 20 years,<sup>28</sup> and the level effect of the equivalence scales applied is comparable to that of the choice of a particular income threshold, say, at 50% or 60%. Choosing the simplified OECD scale could therefore be considered as just another arbitrary choice that is helpful in allowing for rough comparisons across countries as well as over time.

However, if the interest lies with the composition of the population at poverty risk, or with poverty (risks) for specific types of households, equivalence scales ought to become more elaborate. Specifically, their income-dependence matters in these cases to avoid not only substantial measurement error, but also misleading policy conclusions. Fighting poverty requires taking care of the specific situation of different household types affected and cannot be based on aggregate ARP rates. For applications of this kind, the fact that income-dependent scales can only be determined for household types that are observed frequently enough to allow for representative assessments is not much of a limitation.

The task of the present paper is therefore twofold. First, it should be taken as a word of caution. Equivalence scales play a major role in research on income inequality and poverty. Although the (modified) OECD scale has become a quasi-standard in this area, this role should be reflected and probably re-considered. Currently, assumptions underlying this scale are disregarded by many researchers and certainly by the greater public. Yet, blindly following conventional patterns in applied poverty research is not an ideal solution.

Secondly, we would like to stimulate further research in which income-dependent scales should be explored more fully, using different data bases, different approaches and empirical specifications, and they should also be subjected to comparative work at an international level. The goal of such work should be to develop new, more differentiated standards for applied research on distribution and poverty, plus agreements on when and how to use them.

Last but not least, empirically determining equivalence scales that are income-dependent could also inform public policies that are meant to address poverty or to target households of differing size and structure in other ways. For instance, as is highlighted in Muellbauer and van de Ven (2003) and utilized in Bargain et al. (2014) and van de Ven et al. (2017), national tax-benefit schemes rest on implicit equivalence scales, e.g., by the way they are defining entitlements to receive income support for additional household members or the tax treatment of couples and children. It is thus interesting to note that implicit scales for income support and the taxation of low incomes tend to be steeper than the OECD scale. Up to a point, this may reflect national preferences for redistribution. But it might also be useful

---

<sup>28</sup>This is explicitly demonstrated in Garbuszus et al. (2018a, Figure 4).

to design such schemes in the light of more profound knowledge regarding the structure of actual needs at different points in the income distribution.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10888-021-09494-7>.

**Acknowledgements** We would like to thank the Bertelsmann Stiftung for generously supporting our research. We are also grateful to participants of several conferences and workshops for interesting discussions (for instance, at Freie Universität Berlin, the Max Planck Institute for Demographic Research, Rostock, and the German Federal Statistical Office) and to the responsible co-editor as well as two anonymous referees for very helpful comments and suggestions.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Apps, P., Rees, R.: Inequality Measurement and Tax/Transfer Policy. IZA DP 13326 (2020)
- Banks, J., Blundell, R., Lewbel, A.: Quadratic Engel curves and consumer demand. *Rev. Econ. Stat.* **79**(4), 527–539 (1997)
- Bargain, O., Decoster, A., Dolls, M., Neumann, D., Peichl, A., Siegloch, S.: Welfare, labor supply and heterogeneous preferences: evidence for Europe and the US. *Soc. Choice Welf.* **41**(4), 789–817 (2013)
- Bargain, O., Dolls, M., Neumann, D., Peichl, A., Siegloch, S.: Tax-benefit revealed social preferences in Europe and the US. *Ann. Econ. Statist.* **113–114**, 257–289 (2014)
- Biewen, M., Juhasz, A.: Direct estimation of equivalence scales and more evidence on independence of base. *Oxford Bull. Econ. Stat.* **79**(5), 875–905 (2017)
- Blundell, R., Browning, M., Crawford, I.A.: Nonparametric engel curves and revealed preference. *Econometrica* **71**(1), 205–240 (2003)
- Blundell, R., Lewbel, A.: The information content of equivalence scales. *J. Economet.* **50**(1–2), 49–68 (1991)
- Browning, M.: Children and household economic behavior. *J. Econ. Lit.* **30**(3), 1434–1475 (1992)
- Chiappori, P.-A.: Equivalence versus indifference scales. *Econ. J.* **126**(592), 523–545 (2016)
- Cook, R.D.: Influential observations in linear regression. *J. Am. Stat. Ass.* **74**(365), 169–174 (1979)
- Cowell, F.A. *Measuring Inequality*, 3rd edn. Oxford University Press, Oxford (2011)
- Deaton, A.S., Muellbauer, J.: An almost ideal demand system. *Am. Econ. Rev.* **70**(3), 312–326 (1980)
- Donaldson, D., Pendakur, K.: Equivalent-expenditure functions and expenditure-dependent equivalence scales. *J. Pub. Econ.* **88**(1–2), 175–208 (2004)
- Donaldson, D., Pendakur, K.: The identification of fixed costs from consumer behavior. *J. Bus. Econ. Stat.* **24**(3), 255–265 (2006)
- Donaldson, D., Weymark, J.A.: A single-parameter generalization of the gini indices of inequality. *J. Econ. Theory* **22**(1), 67–86 (1980)
- Dudel, C., Garbuszus, J.M., Ott, N., Werding, M.: Matching as non-parametric preprocessing for the estimation of equivalence scales. *J. Econ. Stat.* **237**(2), 115–141 (2017)
- Dudel, C., Garbuszus, J.M., Ott, N., Werding, M.: Income-(in)dependent equivalence scales and inequality measurement. *German Econ. Rev.* (2021)
- Dudel, C., Garbuszus, J.M., Schmied, J.: Assessing differences in household needs: a comparison of approaches for the estimation of equivalence scales using German expenditure data. *Empir. Econ.* **60**(4), 1629–1659 (2020)
- Engel, E.: Die productions- und consumtionsverhältnisse des königreichs sachsen. *Zeitschrift des Statistischen Bureaus des Königlich Sächsischen Ministeriums des Inneren* **3**(8+9), 153–182 (1857)

- Fleurbay, M.: *Fairness, Responsibility and Welfare*. Oxford University Press, Oxford (2002)
- Fuchs, V.R.: Towards a theory of poverty. In: Task Force on Economic Growth and Opportunity (ed.) *The Concept of Poverty*, pp. 71–91. Chamber of Commerce of the United States, Washington (1965)
- Garbuszus, J.M.: *Quadratische Engelkurven. Schätzung von nutzenabhängigen Äquivalenzskalen*, vol. 86 of *Wissenschaftliche Beiträge aus dem Tectum Verlag*. Reihe: Sozialwissenschaften. Tectum Verlag, Baden-Baden (2018)
- Garbuszus, J.M., Ott, N., Pehle, S., Werding, M.: Development of family income since the 1990s: A fresh look at German microdata using income-dependent equivalence scales. SOEPpaper 987 (2018a)
- Garbuszus, J.M., Ott, N., Pehle, S., Werding, M.: Wie hat sich die Einkommenssituation von Familien entwickelt? Bertelsmann Stiftung Gütersloh (2018b)
- Goebel, J., Grabka, M.M., Liebig, S., Kroh, M., Richter, D., Schröder, C., Schupp, J.: The German Socio-Economic Panel Study (SOEP). *Jahrbücher für Nationalökonomie und Statistik* **239**(2), 345–360 (2019)
- Goebel, J., Krause, P., Habich, R.: *Einkommensentwicklung – Verteilung, Angleichung, Armut und Dynamik*. Statistisches Bundesamt & WZB f Sozialforschung (Hrsg.) Datenreport (2016)
- Gower, J.C.: A general coefficient of similarity and some of its properties. *Biometrics*, 857–871 (1971)
- Grabka, M.M., Frick, J.R.: Weiterhin hohes Armutsrisiko in Deutschland: Kinder und junge Erwachsene sind besonders betroffen. *DIW-Wochenbericht* **77**(7), 2–11 (2010)
- Grabka, M.M., Krause, P.: Einkommen und Armut von Familien und älteren Menschen. *DIW-Wochenbericht* **72**(9), 155–162 (2005)
- Hagenaars, A.J., Vos, K., Zaidi, M.A.: Poverty statistics in the late 1980s. Theme / Statistical Office of the European Communities : 3, Population and social conditions : Series C, Accounts, surveys and statistics, Off. of Official Publ. of the Europ. Communities, Luxembourg (1994)
- Hansen, B.B., Klopfer, S.O.: Optimal full matching and related designs via network flows. *J. Comp. Graph. Stat.* **15**(3), 609–627 (2006)
- Howe, H.: Development of the Extended Linear Expenditure System from simple saving assumptions. *Eur. Econ. Rev.* **6**(3), 305–310 (1975)
- Kapteyn, A., Van Praag, B.: A new approach to the construction of family equivalence scales. *Eur. Econ. Rev.* **7**(4), 313–335 (1976)
- Keen, M.: Zero expenditures and the estimation of Engel curves. *J. App. Econ.* **1**(3), 277–286 (1986)
- Klein, L.R., Rubin, H.: A constant-utility index of the cost of living. *Rev. Econ. Stud.* **15**(2), 84–87 (1947)
- Koulouvatianos, C., Schröder, C., Schmidt, U.: On the income dependence of equivalence scales. *J. Pub. Econ.* **89**(5-6), 967–996 (2005)
- Lewbel, A.: Household equivalence scales and welfare comparisons. *J. Public Econ.* **39**(3), 377–391 (1989)
- Lewbel, A., Pendakur, K.: *Equivalence scales*. In: Durlauf, S., Blume, L.E. (eds.) *New Palgrave Dictionary of Economics*. 2nd edn., vol. 3, pp. 26–29. Palgrave MacMillan, Hampshire/New York (2008)
- Lluch, C.: The extended linear expenditure system. *Eur. Econ. Rev.* **4**(1), 21–32 (1973)
- Merz, J.: FELES: The functionalized extended linear expenditure system: Theory, estimation procedures and application to individual household consumption expenditures involving socioeconomic and sociodemographic characteristics. *Eur. Econ. Rev.* **23**(3), 359–394 (1983)
- Muellbauer, J., van de Ven, J.: *Equivalence scales and taxation*. In: Dagum, C., Ferrari, G. (eds.) *Household Behaviour, Equivalence Scales, Welfare and Poverty*, pp. 85–106. Physica, New York (2003)
- Nelson, J.A.: Household equivalence scales: theory versus policy. *J. Lab. Econ.* **11**(3), 471–493 (1993)
- OECD: What are equivalence scales? OECD Income Distribution Database (IDD): Methods and Concepts. <http://www.oecd.org/eo/growth/OECD-Note-EquivalenceScales.pdf>. Last accessed 19 April 2021 (2021)
- Pollak, R.A., Wales, T.J.: *Demand System Specification and Estimation*. Oxford University Press, Oxford (1992)
- Rothbarth, E., Madge, C. (ed.): *Note on a method of determining equivalent income for families of different composition*. Cambridge University Press, Cambridge (1943)
- Statistisches Bundesamt: *Einkommens- und Verbrauchsstichprobe: Aufgabe, Methode und Durchführung*. In: *Wirtschaftsrechnungen*. Statistisches Bundesamt, Wiesbaden (2017)
- Stichnoth, H.: *Verteilungswirkungen ehe- und familienbezogener Leistungen und Maßnahmen*. Heinrich-Böll-Stiftung, Berlin (2016)
- Stone, R.: Linear expenditure systems and demand analysis: an application to the pattern of British demand. *Econ. J.* **64**(255), 511–527 (1954)
- van de Ven, J., Hérault, N., Azpitarte, F.: Identifying tax implicit equivalence scales. *J. Econ. Inequal.* **15**(3), 257–275 (2017)
- Zellner, A.: An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *J. Am. Stat. Ass.* **57**(298), 348–368 (1962)