Economics Bulletin

Volume 35, Issue 4

Nonparametric bounds on equivalence scales

Christian Dudel
Ruhr-Universitaet Bochum

Abstract

Methods for estimating equivalence scales usually rely on rather strong identifying assumptions. This note considers nonparametric bounds on equivalence scales derived from the potential outcomes framework and uses nonparametric methods for estimation, which requires only mild assumptions. The method yields only lower and upper bounds of equivalence scales rather than point estimates. The results of an analysis using German expenditure data show that the range implied by these bounds is rather wide.

Citation: Christian Dudel, (2015) "Nonparametric bounds on equivalence scales", *Economics Bulletin*, Volume 35, Issue 4, pages 2161-2165
Contact: Christian Dudel - christian.dudel@rub.de.
Submitted: July 25, 2015. **Published:** October 16, 2015.

1 Introduction

Household equivalence scales are routinely applied in research on poverty and inequality. They are used to adjust household income (or expenditure) of households of different size and composition. The resulting equivalized income is assumed to be directly comparable across households. More specifically, equivalence scales indicate how much more income a household of type a needs to reach the same welfare level as a reference household of type b.

This note builds on the potential outcomes framework (e.g. Holland, 1986) and suggests nonparametric bounds for equivalence scales that depend only on mild assumptions and, following a suggestion by Szulc (2009), can be used with any set of welfare indicators. In addition, estimation procedures are nonparametric compared with approaches found in the literature, which are usually parametric or semiparametric. These approaches achieve identification using the assumption that equivalence scales do not depend on the welfare level, or they assume a certain parametric structure of cost or (indirect) utility functions, or they require both assumptions. The potential outcomes approach requires neither independence of the welfare level nor assumptions on the functional form.

The remainder of this note is structured as follows. The potential outcomes framework is introduced in section 2. Nonparametric bounds and nonparametric estimation are described in section 3. An empirical example is presented in section 4 and the results are compared to equivalence scales found in the literature. Section 5 concludes.

2 Potential outcomes and equivalence scales

Let C be an indicator variable that equals 1 if a specific treatment has been received and 0 otherwise; Y denotes an outcome of interest. The basic reasoning of the potential outcomes framework is that for each unit i of a sample of n units there are two potential outcomes: y_i^1 is the outcome given the treatment and y_i^0 is the outcome that would be realized in the absence of the treatment. In many cases, analysis is concerned with the average treatment effect (ATE) $\mathrm{E}(Y^1 - Y^0)$.

Because in practice either y_i^1 or y_i^0 can be observed and never both, certain assumptions are needed to estimate the ATE (e.g. Imbens, 2004). The first condition that needs to be met is unconfoundedness, which requires that the pair Y^1, Y^0 is independent of C given a vector of variables $\mathbf{X}$. The second condition requires that $0 < \Pr(C = 1|\mathbf{X}) < 1$ (overlap condition). Additionally, the stable unit treatment value assumption is invoked, which implies the independence of the outcomes of observation i from the treatment status of observation j.

Adapting ideas to the case of household equivalence scales, for each household i and each of two possible household compositions $C = 0$ (e.g. childless couple) and $C = 1$ (e.g. couple with one child), there are two pairs of potential outcomes: u_i^0 and $y_i^0(u_i^0)$ are the level of welfare respectively the income needed to reach this level given composition $c_i = 0$; u_i^1 and $y_i^1(u_i^1)$ are the welfare level and the income needed to reach this welfare level given composition $c_i = 1$.

Household welfare U is assumed to be a function of $\mathbf{X}$, which is a vector of household characteristics that capture household welfare. For example, $\mathbf{X}$ could include the expenditure share on food (Engel approach) or the expenditure on some good that is only consumed by adults (Rothbarth approach). Furthermore, $\mathbf{X}$ could include additional characteristics like the age or education of household members. To keep the notation simple, $\mathrm{E}(Y^C|U)$

will be used instead of $\mathrm{E}(Y^C(U)|U)$. Moreover, note that $\mathrm{E}(\mathrm{E}(Y^C|U)) = \mathrm{E}(\mathrm{E}(Y^C|\mathbf{X}))$.

A household-specific equivalence scale can then be defined as the ratio $y_i^1(u)/y_i^0(u)$ for some welfare level u. Note that this phrases equivalence scales in slightly different terms than is usually done in the literature. In microeconomic theory equivalence scales are defined as the ratio of cost functions $\mathcal{K}(\mathbf{p}, u, C = 1)/\mathcal{K}(\mathbf{p}, u, C = 0)$, where $\mathcal{K}(\cdot)$ denotes the cost function, $\mathbf{p}$ is a vector of prices, u is the utility level, and C indicates household composition. Thus, as outlined in the introduction, equivalence scales show how much more income a household consisting of a couple with one child needs to reach the same utility level as a childless couple. The potential outcomes approach, on the other hand, starts from the question how much more income household i needs if household composition changes from $C = 0$ to $C = 1$, i.e. if a child enters the household. Somewhat simplified the former approach compares groups a and b, while the latter assesses the impact of changing from group a to group b. Depending on the question a researcher wants to answer one approach may be more suited than the other. For example, if one is interested in giving policy advice on the amount of child benefits the potential outcomes framework may be more appropriate, as it directly tries to tackle the question how welfare and needs change.

For estimation of the potential outcomes approach again only one of the pairs of potential outcomes is observed for each unit i. That is, only $y_i = c_i y_i^1(u_i^1) + [1 - c_i] y_i^0(u_i^0)$ and $u_i = c_i u_i^1 + [1 - c_i] u_i^0$ are known. Furthermore, even if both pairs of outcomes could be observed, this would not suffice to calculate the ratio given above because only $y_i^1(u_i^1)/y_i^0(u_i^0)$ could be calculated and not $y_i^1(u_i^0)/y_i^0(u_i^0)$ or $y_i^1(u_i^1)/y_i^0(u_i^1)$. The reasoning behind the proposal of Szulc (2009) is that for households of composition $C = 0$, $Y^0(U^0) = Y^0(\mathbf{X})$ is known. If U is a function of an observed set of welfare indicators $\mathbf{X}$, $Y^1(U^0) = Y^1(\mathbf{X})$ can be estimated using households of composition $C = 1$ and the same value of $\mathbf{X}$. For households of composition $C = 1$, $Y^1(U^1) = Y^1(\mathbf{X})$ is known and $Y^0(\mathbf{X})$ can be estimated from households of composition $C = 0$.

3 Nonparametric bounds on equivalence scales

Given the three assumptions introduced in the preceding section the joint distribution of Y^0 and Y^1 is not identified (e.g. Abbring and Heckman, 2007). Only the marginal distributions are known. Because of this, the expected value and the distribution of the ratio Y^1/Y^0 are not identified. More formally, $\mathrm{E}(Y^1/Y^0)$ can be written as

$$\mathrm{E}\left[\mathrm{E}\left(\frac{Y^1}{Y^0}\middle|\mathbf{X}\right)\right] = \mathrm{E}\left[\frac{\mathrm{E}(Y^1|\mathbf{X})}{\mathrm{E}(Y^0|\mathbf{X})}\right] - \mathrm{E}\left[\frac{1}{\mathrm{E}(Y^0|\mathbf{X})}\mathrm{Cov}\left(\frac{Y^1}{Y^0}, Y^0\middle|\mathbf{X}\right)\right]. \tag{1}$$

Given unconfoundedness, the conditional expectations $\mathrm{E}(Y^1|\mathbf{X})$ and $\mathrm{E}(Y^0|\mathbf{X})$ on the right-hand side can easily be estimated, but the assumptions stated in the preceding section do not suffice to estimate the covariance of Y^1/Y^0 and Y^0.

Fréchet (1951) established that the marginal distributions of Y^1 and Y^0 can be used to derive bounds on the joint distribution. Using this result and following Cambanis et al. (1976), Fan and Zhu (2009) showed that sharp nonparametric bounds on $\mathrm{E}[k(y^1, y^0)|\mathbf{X}]$ can be obtained from these upper and lower bounds where $k(y^1, y^0)$ is a strictly subadditive (strictly quasi-antitone) function of y^1 and y^0. Let $\beta^L(\mathbf{X})$ and $\beta^U(\mathbf{X})$ denote the lower and upper bound of $\mathrm{E}[k(y^1, y^0)|\mathbf{X}]$, respectively. These bounds can be calculated by

$$\beta^L(\mathbf{X}) = \int_0^1 k\left(F_1^{-1}(t|\mathbf{X}), F_0^{-1}(t|\mathbf{X})\right) dt \tag{2}$$

and

$$\beta^U(\mathbf{X}) = \int_0^1 k\left(F_1^{-1}(t|\mathbf{X}), F_0^{-1}(1-t|\mathbf{X})\right) dt, \tag{3}$$

where $F_1^{-1}(u|\mathbf{X})$ and $F_0^{-1}(u|\mathbf{X})$ are the quantile functions of the conditional marginal distributions of Y^1 and Y^0, respectively. If k is a superadditive function, the bounds are reversed, so that (2) gives the upper and (3) the lower bound.

Note that the results of Cambanis et al. (1976) only require k to be strictly subadditive or strictly superadditive on the support of Y^1 and Y^0. If we assume that $Y^1 > 0$ and $Y^0 > 0$, which is reasonable for income, $k(y^1, y^0) = y^1/y^0$ is subadditive and sharp bounds of the conditional expectation $\mathrm{E}(Y^1/Y^0|\mathbf{X})$ can be calculated using (2) and (3). Furthermore, these bounds can be used to derive bounds on the conditional covariance, which are given by

$$\begin{aligned}\mathrm{E}(Y^1|\mathbf{X}) - \mathrm{E}^U(Y^1/Y^0|\mathbf{X})\mathrm{E}(Y^0|\mathbf{X}) \leq \\ \mathrm{Cov}(Y^1/Y^0, Y^0|\mathbf{X}) \leq \mathrm{E}(Y^1|\mathbf{X}) - \mathrm{E}^L(Y^1/Y^0|\mathbf{X})\mathrm{E}(Y^0|\mathbf{X}),\end{aligned} \tag{4}$$

where $\mathrm{E}^L(Y^1/Y^0|\mathbf{X})$ and $\mathrm{E}^U(Y^1/Y^0|\mathbf{X})$ are the lower and upper bound of the conditional expectation, respectively. As the bounds on $\mathrm{E}(Y^1/Y^0|\mathbf{X})$ are sharp, the bounds in (4) are also sharp. This allows for the calculation of bounds of the second term on the right-hand side of (1), which can be interpreted as a correction term, while the first term can be seen as a naive estimate. The latter is an interesting benchmark for the results of other methods and can be seen as a static group comparison, while the correction term accounts for the fact that the starting point of analysis is a change from one group to the other, as outlined in the previous section.

Let β^L and β^U denote the lower and upper bound of the unconditional expectation $\mathrm{E}(Y^1/Y^0)$. A plug-in estimator is given by

$$\hat{\beta}^L = \frac{1}{n}\sum_{i=1}^{n} \hat{\beta}^L(\mathbf{x}_i) \quad \text{and} \quad \hat{\beta}^U = \frac{1}{n}\sum_{i=1}^{n} \hat{\beta}^U(\mathbf{x}_i), \tag{5}$$

where $\hat{\beta}^L(\mathbf{x}_i)$ and $\hat{\beta}^U(\mathbf{x}_i)$ are estimates of (2) and (3) using $k(y^1, y^0) = y^1/y^0$, respectively (Fan and Zhu, 2009). To estimate (2) and (3), kernel estimators for the conditional cumulative distribution functions $F_C(u|\mathbf{x}_i)$ are used, as proposed by Fan and Zhu (2009). Estimates of $F_C(u|\mathbf{x}_i)$ can also be used to estimate $\mathrm{E}(Y^C|\mathbf{x}_i)$, which in combination with $\hat{\beta}^L(\mathbf{x}_i)$ and $\hat{\beta}^U(\mathbf{x}_i)$ allows the calculation of bounds on $\mathrm{Cov}(Y^1/Y^0, Y^0|\mathbf{x}_i)$. Finally, the latter estimates can be used to calculate the first term and bounds on the second term on the right-hand side of (1), i.e. the naive estimate and the correction term.

Note that the procedure outlined above does not require equivalence scales to be independent of the welfare level, the so called base independence assumption (Pendakur, 1999). This assumption implies that $\mathrm{E}\left(Y^1/Y^0|\mathbf{X}\right) = \mathrm{E}\left(Y^1/Y^0\right)$. If base independence is violated, the estimator can still be applied and gives bounds on the expected value of equivalence weights over all welfare levels.

Table 1: Estimates of nonparametric bounds on equivalence scales for couples with one child and childless couples as reference

Welfare indicator(s)	Lower bound	Upper bound	Naive estimate
Expenditure share food	1.11	1.43	1.10
Expenditure clothing/adults	1.04	1.45	1.03
Housing	1.12	1.50	1.10
All	1.16	1.46	1.14

4 Empirical example

The estimator was applied to data from the German Income and Expenditure Survey 2008 ("*Einkommens- und Verbrauchsstichprobe*", EVS). Equivalence scales were estimated for couples with one child (less than 14 years of age) using childless couples as a reference and net household income as the outcome. The following welfare indicators were used: the expenditure share for food (Engel method); the expenditure on clothing for adults (Rothbarth method); and the homeownership status and housing space per household member (housing). All of the welfare indicators were used separately and in combination. Both the Engel method and the Rothbarth method are well known approaches to assess the welfare level of households (e.g. Deaton and Muellbauer, 1986). Housing follows a suggestion of Szulc (2009).

Additional household characteristics included in the analysis cover the following: the age of the household head; the education of the household head; the region of the household (West or East Germany); and whether both partners are employed (dual-earner household). Analysis was carried out with data on 7116 childless couples and 2249 couples with one child. Calculations were done using the freely available statistical package R and the methods applied for nonparametric estimation and bandwidth selection were as included in the `np` package by Hayfield and Racine (2008). R code is available upon request.

The results are shown in Table 1, which also includes naive estimates calculated using only the first term on the right-hand side of (1). The bounds are rather wide. For instance, the difference between the upper and lower bound amounts to 0.32 in the case of the expenditure share of food. If one assumes that 1.5 is a priori a plausible upper bound for the equivalence scale for couples with children compared to childless couples, the upper bounds in table 1 are not informative in the case of housing. Because both the lower and upper bounds of the covariance of Y^1/Y^0 and Y^0 are negative, the lower bounds shown in table 1 are close to the naive estimates that only make use of the first term on the right-hand side of (1). Stated differently, the results of the naive approach lie outside the identification bounds. This shows that the two perspectives outlined in section 2 lead to quite different results and a static group comparison will underestimate the effect of adding a child to a household.

Results for equivalence scales for Germany taken from the literature were generally derived using parametric methods and stronger identifying assumptions. Most are close to the lower bounds and naive estimates shown in table 1. For example, Faik (2011) and Schulte (2007) both used expenditure data and arrived at estimates of 1.08 and 1.21, respectively. Another example are the results given by Schwarze (2003), which range from 1.10 to 1.15 and make use of data on satisfaction with household income. As these authors apply methods based on group comparisons this comes as no surprise.

5 Concluding remarks

In this paper a method for the estimation of nonparametric bounds on equivalence scales was proposed that relies on nonparametric estimators and only requires rather mild assumptions. The approach can be used with any combination of welfare indicators. For instance, satisfaction with the financial situation of the household could be used in addition to, or instead of, expenditure data. An example shows that the range of possible values, as indicated by the lower and upper bound, tends to be large.

References

Abbring, J. H., and Heckman, J. J. (2007). "Econometric evaluation of social programs, Part III." In *Handbook of Econometrics, Volume 6B*, 5146–5303, Elsevier.

Cambanis, S., Simons, G., and Stout, W. (1976). "Inequalitites for $\mathrm{E}k(x, y)$ when the marginals are fixed." *Probability Theory and Related Fields*, *36*, 285–294.

Deaton, A., and Muellbauer, J. (1986). "On measuring child costs: With application to poor countries." *Journal of Political Economy*, *94*, 720–744.

Faik, J. (2011). "Der Zerlegungs-Ansatz – ein alternativer Vorschlag zur Messung von Armut." *AStA – Wirtschafts- und Sozialstatistisches Archiv*, *4*, 293–315.

Fan, Y., and Zhu, D. (2009). "Partial identification and confidence sets for functionals of the joint distribution of potential outcomes.", Working Paper, Vanderbilt University.

Fréchet, M. (1951). "Sur les tableaux de corrélation dont les marges sont données." *Annales de l'Université de Lyon, Sciences*, *9*, 53–77.

Hayfield, T., and Racine, J. S. (2008). "Nonparametric econometrics: The np package." *Journal of Statistical Software*, *27*, 1–32.

Holland, P. (1986). "Statistics and causal inference." *Journal of the American Statistical Association*, *81*, 945–960.

Imbens, G. W. (2004). "Nonparametric estimation of average treatment effects under exogeneity: A review." *Review of Economics and Statistics*, *86*, 4–29.

Pendakur, K. (1999). "Semiparametric estimates and tests of base-independent equivalence scales." *Journal of Econometrics*, *88*, 1–40.

Schulte, J. (2007). *Equivalence Scales: Identication and Estimation. A Cross-Sectional Analysis of German Data.* Ph.D. thesis, Freie Universität Berlin.

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

Szulc, A. (2009). "A matching estimator of household equivalence scales." *Economics Letters*, *103*, 81–83.