

Biewen, Martin; Jenkins, Stephen P.

Working Paper

Estimation of Generalized Entropy and Atkinson inequality indices from survey data

ISER Working Paper Series, No. 2003-11

Provided in Cooperation with:

Institute for Social and Economic Research (ISER), University of Essex

Suggested Citation: Biewen, Martin; Jenkins, Stephen P. (2003) : Estimation of Generalized Entropy and Atkinson inequality indices from survey data, ISER Working Paper Series, No. 2003-11, University of Essex, Institute for Social and Economic Research (ISER), Colchester

This Version is available at:

<https://hdl.handle.net/10419/92108>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



ESTIMATION OF GENERALIZED ENTROPY AND ATKINSON INEQUALITY INDICES FROM SURVEY DATA

**Martin Biewen
(University of Mannheim)**

and

**Stephen P. Jenkins
(ISER, University of Essex)**

**ISER Working Papers
Number 2003-11**

Institute for Social and Economic Research

The Institute for Social and Economic Research (ISER) specialises in the production and analysis of longitudinal data. ISER incorporates the following centres:

- ESRC Research Centre on Micro-social Change. Established in 1989 to identify, explain, model and forecast social change in Britain at the individual and household level, the Centre specialises in research using longitudinal data.
- ESRC UK Longitudinal Studies Centre. This national resource centre was established in October 1999 to promote the use of longitudinal data and to develop a strategy for the future of large-scale longitudinal surveys. It was responsible for the British Household Panel Survey (BHPS) and for the ESRC's interest in the National Child Development Study and the 1970 British Cohort Study
- European Centre for Analysis in the Social Sciences. ECASS is an interdisciplinary research centre which hosts major research programmes and helps researchers from the EU gain access to longitudinal data and cross-national datasets from all over Europe.

The British Household Panel Survey is one of the main instruments for measuring social change in Britain. The BHPS comprises a nationally representative sample of around 5,500 households and over 10,000 individuals who are reinterviewed each year. The questionnaire includes a constant core of items accompanied by a variable component in order to provide for the collection of initial conditions data and to allow for the subsequent inclusion of emerging research and policy concerns.

Among the main projects in ISER's research programme are: the labour market and the division of domestic responsibilities; changes in families and households; modelling households' labour force behaviour; wealth, well-being and socio-economic structure; resource distribution in the household; and modelling techniques and survey methodology.

BHPS data provide the academic community, policymakers and private sector with a unique national resource and allow for comparative research with similar studies in Europe, the United States and Canada.

BHPS data are available from the Data Archive at the University of Essex
<http://www.data-archive.ac.uk>

Further information about the BHPS and other longitudinal surveys can be obtained by telephoning +44 (0) 1206 873543.

The support of both the Economic and Social Research Council (ESRC) and the University of Essex is gratefully acknowledged. The work reported in this paper is part of the scientific programme of the Institute for Social and Economic Research.

Acknowledgement: Financial support from the Deutsche Forschungsgesellschaft (DFG) and core funding to ISER from the ESRC and the University of Essex is gratefully acknowledged.

Readers wishing to cite this document are asked to use the following form of words:

Biewen, M. and Jenkins, Stephen P. (April 2003) 'Estimation of Generalized Entropy and Atkinson inequality indices from survey data', *Working Papers of the Institute for Social and Economic Research*, paper 2003–11. Colchester: University of Essex.

For an on-line version of this working paper and others in the series, please visit the Institute's website at: <http://www.iser.essex.ac.uk/pubs/workpaps/>

Institute for Social and Economic Research
University of Essex
Wivenhoe Park
Colchester
Essex
CO4 3SQ UK
Telephone: +44 (0) 1206 872957
Fax: +44 (0) 1206 873151
E-mail: iser@essex.ac.uk
Website: <http://www.iser.essex.ac.uk>

© April 2003

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted, in any form, or by any means, mechanical, photocopying, recording or otherwise, without the prior permission of the Communications Manager, Institute for Social and Economic Research.

ABSTRACT

Applying a method suggested by Woodruff (*Journal of the American Statistical Association*, 1971), we derive the sampling variances of Generalized Entropy and Atkinson inequality indices when estimated from complex survey data. It turns out that this method also greatly simplifies the calculations for the i.i.d. case when compared to previous derivations in the literature. Both cases are illustrated with examples from the German Socio-Economic Panel Study and the British Household Panel Survey.

NON-TECHNICAL SUMMARY

Probability weighting, clustering, and stratification, are survey design features underlying much of the survey data that economists and others use. It is well known that these features have a potentially large impact on the sampling variability of statistics computed from such surveys. Nevertheless, they are rarely taken into account in practical work, the measurement of inequality being no exception. We derive estimates for the sampling variance of two commonly-used classes of inequality indices, the Generalized Entropy and the Atkinson families, using the approach of Woodruff (*Journal of the American Statistical Association*, 1971). It turns out that Woodruff's method also greatly simplifies the computation of variance estimates in an i.i.d. framework when compared to previous derivations in the literature.

In order to assess the error made by not taking into account clustering and stratification, we apply our results to samples extracted from the German Socio-Economic Panel (GSOEP) and the British Household Panel Survey (BHPS), calculating various estimates of the inequality of income among individuals.

The results suggest that ignoring stratification and first-stage clustering has only a small impact on the sampling variances of indices of income inequality. (Accounting for the clustering of individuals within households appears much more important.) These empirical findings might be interpreted as good news for practitioners using these data, or for those using surveys in which primary sampling unit and strata identification variables are not made available, but it is not clear that these findings may be generalized. Whatever the case, our variance estimators provide a straightforward means by which researchers can accommodate a range of design effects in their analysis.

Estimation of Generalized Entropy and Atkinson Inequality Indices from Complex Survey Data¹

Martin Biewen

University of Mannheim

IZA, Bonn

DIW Berlin

Stephen P. Jenkins

ISER, University of Essex

IZA, Bonn

DIW Berlin

This version: 11 April 2003

Abstract. Applying a method suggested by Woodruff (1971), we derive the sampling variances of Generalized Entropy and Atkinson inequality indices when estimated from complex survey data. It turns out that this method also greatly simplifies the calculations for the i.i.d. case when compared to previous derivations in the literature. Both cases are illustrated with examples from the German Socio-Economic Panel Study and the British Household Panel Survey.

JEL-Classification: C14, D31

Keywords: Inequality, Statistical Inference, Complex Surveys

Correspondence:

Martin Biewen, Department of Economics, University of Mannheim, Verfügungsgebäude L7, 3-5, 68131 Mannheim, Germany, Fax: +49-621-1811841, biewen@rhein.vwl.uni-mannheim.de

¹Financial support from the the Deutsche Forschungsgesellschaft (DFG) and core funding to ISER from the ESRC and the University of Essex is gratefully acknowledged.

1 Introduction

Probability weighting, clustering, and stratification, are survey design features underlying much of the survey data that economists and others use. It is well known that these features have a potentially large impact on the sampling variability of statistics computed from such surveys. Nevertheless, they are rarely taken into account in practical work, the measurement of inequality being no exception. We derive estimates for the sampling variance of two commonly-used classes of inequality indices, the Generalized Entropy and the Atkinson family of indices, using the approach of Woodruff (1971).² It turns out that Woodruff's method also greatly simplifies the computation of variance estimates in an i.i.d. framework when compared to previous derivations in the literature. In order to assess the error made by not taking into account clustering and stratification we apply our results to a sample extracted from the German Socio-Economic Panel (GSOEP) and the British Household Panel Survey (BHPS).

2 Estimation from complex surveys

Generalized Entropy and Atkinson inequality measures can be written, as we show below, as functions $I = f(T)$ of population totals $T = (T_1, \dots, T_K)$. Population total T_k , $k = 1, \dots, K$ is given by the summation of an observational variable t_{hijk} over the different stages of the sampling design, i.e. $T_k = \sum_{n=1}^L \sum_{i=1}^{N_h} \sum_{j=1}^{M_i} t_{hijk}$ where L denotes the number of strata, N_h the number of clusters in stratum h and M_i the number of individuals in cluster i . If the sampling design involves more than one stage of clustering it suffices to consider the first stage only (see e.g. Cochran, 1977). Replacing totals T by their estimates \hat{T} , the index is then estimated as $\hat{I} = f(\hat{T})$ with $\hat{T}_k = \sum_{n=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_i} w_{hij} t_{hijk}$ where n_h is the number of actually sampled first stage clusters and m_i the number of actually sampled individuals in cluster i . The sampling weight of individual hij is given by w_{hij} . Assuming that the sample is large enough that a Taylor approximation of $f(\cdot)$ holds, the variance of \hat{I} can be approximated by the variance of the first order residual $\sum_{k=1}^K (\partial f(T)/\partial T_k) \hat{T}_k$. Woodruff (1971) observed that this variance can be easily determined by reversing the order of summation in the residual, i.e. $\text{var}(\hat{I}) \approx \text{var}(\sum_{n=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_i} w_{hij} [\sum_{k=1}^K (\partial f(T)/\partial T_k) t_{hijk}]) = \text{var}(\hat{S})$. Note that \hat{S} is of the

²An alternative but conceptually less straightforward approach to variance estimation in complex surveys is the estimating equations approach described in Binder (1983) and Binder and Patak (1994). It turns out that this approach leads to the same estimators derived here. Calculations are available from the authors on request.

same form as the \hat{T}_k s so that the problem is reduced to the estimation of the sampling variance of a total estimator for which well-known formulas exist (see Cochran, 1977, or Deaton, 1997). Using these formulas (and replacing T by \hat{T} in the derivative), the variance estimate for \hat{I} is

$$\widehat{\text{var}}(\hat{I}) = \sum_{n=1}^L \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} \left(\sum_{j=1}^{m_i} w_{hij} \tilde{s}_{hij} - \frac{\sum_{i=1}^{n_h} \sum_{j=1}^{m_i} w_{hij} \tilde{s}_{hij}}{n_h} \right)^2 \quad (1)$$

with $\tilde{s}_{hij} = \sum_{k=1}^K (\partial f(\hat{T}) / \partial \hat{T}_k) t_{hijk}$.

If y_{hij} represents the income of individual hij , then population Generalized Entropy and Atkinson indices are given by

$$I_{GE}^\alpha = (\alpha^2 - \alpha)^{-1} \left[U_0^{\alpha-1} U_1^{-\alpha} U_\alpha - 1 \right], \quad \alpha \in \mathbb{R} \setminus \{0, 1\} \quad (2)$$

$$I_{Theil} = T_{1,1} U_1^{-1} - \log(U_1 U_0^{-1}), \quad \alpha \rightarrow 1 \quad (3)$$

$$I_{MLD} = -T_{0,1} U_0^{-1} + \log(U_1 U_0^{-1}), \quad \alpha \rightarrow 0 \quad (4)$$

$$I_A^\varepsilon = 1 - U_0^{-\varepsilon/(1-\varepsilon)} U_1^{-1} U_{1-\varepsilon}^{1/(1-\varepsilon)}, \quad \varepsilon \geq 0, \quad \varepsilon \neq 1, \quad (5)$$

$$I_A^1 = 1 - U_0 U_1^{-1} \exp(T_{0,1} U_0^{-1}), \quad \varepsilon \rightarrow 1 \quad (6)$$

for totals $U_\alpha = \sum_{n=1}^L \sum_{i=1}^{N_h} \sum_{j=1}^{M_i} (y_{hij})^\alpha$ and $T_\alpha = \sum_{n=1}^L \sum_{i=1}^{N_h} \sum_{j=1}^{M_i} (y_{hij})^\alpha (\log y_{hij})$. Note that U_0 is the population size. Estimates of these indices are obtained by replacing U_α , T_α with $\hat{U}_\alpha = \sum_{n=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_i} w_{hij} (y_{hij})^\alpha$ and $\hat{T}_\alpha = \sum_{n=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_i} w_{hij} (y_{hij})^\alpha (\log y_{hij})$.

The corresponding variances are then given by substituting \tilde{s}_{hij} in (1) with

$$\tilde{s}_{hij}^{GE} = \frac{1}{\alpha} \hat{U}_\alpha \hat{U}_1^{-\alpha} \hat{U}_0^{\alpha-2} - \frac{1}{\alpha-1} \hat{U}_\alpha \hat{U}_1^{-\alpha-1} \hat{U}_0^{\alpha-1} \cdot y_{hij} + \frac{1}{\alpha^2 - \alpha} \hat{U}_0^{\alpha-1} \hat{U}_1^{-\alpha} \cdot (y_{hij})^\alpha \quad (7)$$

$$\tilde{s}_{hij}^{Theil} = \hat{U}_1^{-1} \cdot y_{hij} \log y_{hij} - \hat{U}_1^{-1} (\hat{T}_{1,1} \hat{U}_1^{-1} + 1) \cdot y_{hij} + \hat{U}_0^{-1} \quad (8)$$

$$\tilde{s}_{hij}^{MLD} = -\hat{U}_0^{-1} \cdot \log y_{hij} + \hat{U}_1^{-1} \cdot y_{hij} + U_0^{-1} (\hat{T}_{0,1} \hat{U}_0^{-1} - 1) \quad (9)$$

$$\tilde{s}_{hij}^A = \frac{\varepsilon}{1-\varepsilon} \hat{U}_1^{-1} \hat{U}_{1-\varepsilon}^{\frac{1}{1-\varepsilon}} \hat{U}_0^{-\frac{1}{1-\varepsilon}} + \hat{U}_0^{\frac{-\varepsilon}{1-\varepsilon}} \hat{U}_{1-\varepsilon}^{\frac{1}{1-\varepsilon}} \hat{U}_1^{-2} \cdot y_{hij} - \frac{1}{1-\varepsilon} \hat{U}_0^{\frac{-\varepsilon}{1-\varepsilon}} \hat{U}_1^{-1} \hat{U}_{1-\varepsilon}^{\frac{\varepsilon}{1-\varepsilon}} \cdot (y_{hij})^{1-\varepsilon} \quad (10)$$

$$\tilde{s}_{hij}^1 = (\hat{I}_A^1 - 1) \hat{U}_0^{-1} (1 - \hat{U}_0^{-1} \hat{T}_{0,1}) + (1 - \hat{I}_A^1) \hat{U}_1^{-1} \cdot y_{hij} + (\hat{I}_A^1 - 1) \hat{U}_0^{-1} \cdot \log y_{hij}. \quad (11)$$

These variance estimators allow arbitrary correlations between observations below the first-stage clusters. They are therefore an alternative to the estimators developed by Schluter and Trede (2002).

3 Application to the i.i.d. case

In an i.i.d. framework the above indices are usually treated as follows. Income x_i and weight w_i of observation $i = 1, \dots, n$ are regarded as i.i.d. draws from a population (x, w) .³ The index in question can then be represented as a function $I = g(\mu)$ of population moments $\mu = E(X_i)$, where X_i is a vector-valued function of (x_i, w_i) . It is estimated as $\hat{I} = g(\bar{X})$ and its sampling variance as $n^{-1} \nabla g(\bar{X})' \widehat{\text{var}}(X_i) \nabla g(\bar{X})$ (Cowell, 1989, or for the case without weights, Thistle, 1990). By contrast, Woodruff's method would yield a variance estimate $n^{-1} \widehat{\text{var}}(\nabla g(\bar{X})' X_i)$. It is easy to see that both estimates are identical. However, Woodruff's method leads to much simpler expressions as the problem is reduced to estimating the sampling variance of a scalar. In particular, no covariances need to be computed. Defining $\mu_\alpha = E(w_i x_i^\alpha)$, $\tau_\alpha = E(w_i x_i^\alpha (\log x_i))$, $\hat{\mu}_\alpha = n^{-1} \sum_{i=1}^n w_i x_i^\alpha$ and $\hat{\tau}_{\alpha,\gamma} = n^{-1} \sum_{i=1}^n w_i x_i^\alpha (\log x_i)$, the sampling variances of the indices

$$I_{GE}^\alpha = (\alpha^2 - \alpha)^{-1} [\mu_0^{\alpha-1} \mu_1^{-\alpha} \mu_\alpha - 1], \quad \alpha \in \mathbb{R} \setminus \{0, 1\} \quad (12)$$

$$I_{Theil} = \tau_{1,1} \mu_1^{-1} - \log(\mu_1 \mu_0^{-1}), \quad \alpha \rightarrow 1 \quad (13)$$

$$I_{MLD} = -\tau_{0,1} \mu_0^{-1} + \log(\mu_1 \mu_0^{-1}), \quad \alpha \rightarrow 0 \quad (14)$$

$$I_A^\varepsilon = 1 - \mu_0^{-\varepsilon/(1-\varepsilon)} \mu_1^{-1} \mu_{1-\varepsilon}^{1/(1-\varepsilon)}, \quad \varepsilon \geq 0, \quad \varepsilon \neq 1, \quad (15)$$

$$I_A^1 = 1 - \mu_0 \mu_1^{-1} \exp(\tau_{0,1} \mu_0^{-1}), \quad \varepsilon \rightarrow 1 \quad (16)$$

can therefore simply be estimated by substituting

$$\tilde{z}_i^{GE} = \frac{1}{\alpha} \hat{\mu}_\alpha \hat{\mu}_1^{-\alpha} \hat{\mu}_0^{\alpha-2} - \frac{1}{\alpha-1} \hat{\mu}_\alpha \hat{\mu}_1^{-\alpha-1} \hat{\mu}_0^{\alpha-1} \cdot x_i + \frac{1}{\alpha^2 - \alpha} \hat{\mu}_0^{\alpha-1} \hat{\mu}_1^{-\alpha} \cdot x_i^\alpha \quad (17)$$

$$\tilde{z}_i^{Theil} = \hat{\mu}_1^{-1} \cdot x_i \log x_i - \hat{\mu}_1^{-1} (\hat{\tau}_{1,1} \hat{\mu}_1^{-1} + 1) \cdot x_i + \hat{\mu}_0^{-1} \quad (18)$$

$$\tilde{z}_i^{MLD} = -\hat{\mu}_0^{-1} \cdot \log x_i + \hat{\mu}_1^{-1} \cdot x_i + \mu_0^{-1} (\hat{\tau}_{0,1} \hat{\mu}_0^{-1} - 1) \quad (19)$$

$$\tilde{z}_i^A = \frac{\varepsilon}{1-\varepsilon} \hat{\mu}_1^{-1} \hat{\mu}_{1-\varepsilon}^{\frac{1}{1-\varepsilon}} \hat{\mu}_0^{-\frac{1}{1-\varepsilon}} + \hat{\mu}_0^{\frac{-\varepsilon}{1-\varepsilon}} \hat{\mu}_{1-\varepsilon}^{\frac{1}{1-\varepsilon}} \hat{\mu}_1^{-2} \cdot x_i - \frac{1}{1-\varepsilon} \hat{\mu}_0^{\frac{-\varepsilon}{1-\varepsilon}} \hat{\mu}_1^{-1} \hat{\mu}_{1-\varepsilon}^{\frac{\varepsilon}{1-\varepsilon}} \cdot x_i^{1-\varepsilon} \quad (20)$$

$$\tilde{z}_i^1 = (\hat{I}_A^1 - 1) \hat{\mu}_0^{-1} (1 - \hat{\mu}_0^{-1} \hat{\tau}_{0,1}) + (1 - \hat{I}_A^1) \hat{\mu}_1^{-1} \cdot x_i + (\hat{I}_A^1 - 1) \hat{\mu}_0^{-1} \cdot \log x_i \quad (21)$$

for \tilde{z}_i in

$$\widehat{\text{var}}(\hat{I}) = \frac{1}{n(n-1)} \sum_{i=1}^n \left(w_i \tilde{z}_i - \frac{\sum_{i=1}^n w_i \tilde{z}_i}{n} \right)^2. \quad (22)$$

Equation (22) has a similar structure to equation (1), but note that the weights are treated differently in the complex survey and i.i.d. cases.⁴

³If the distribution of household income among individuals is analyzed, then observational units i are households and observations are replicated at the household level. In this case weights are $w_i = w'_i w''_i$ with household size w'_i and sample weight w''_i . See Biewen (2002).

⁴For more discussion of different weighting concepts, see Cowell and Jenkins (2003).

4 Empirical illustration

We contrast the variance estimators using data from the first waves of the German Socio-Economic Panel (SOEP Group, 2001), and the British Household Panel Survey (Taylor et al., 2002).⁵ The GSOEP and the BHPS are widely used to analyze the income distribution in these two countries. Moreover, they provide information on primary sampling units (clusters) and strata identification variables. The first wave of each survey was chosen to avoid complications of the panel design. We considered the distribution of income among individuals. Following convention, each person was assumed to receive the equivalent household income of the household to which she belonged. (The equivalence scale was the square root of household size.)

The survey estimates shown in column (1) of Table 1 take into account sampling weights, stratification and clustering. Replication of observations at the household level is automatically accounted for, as this represents a form of clustering below the first stage clusters. The i.i.d. estimates shown in column (2) only take into account sampling weights and replication of observations at the household level, but not stratification or clustering. A comparison of columns (1) and (2) indicates that ignoring clustering and stratification makes surprisingly little difference for these data sets. By contrast, the estimates shown in columns (3) and (4) suggest that taking into account the replication of observations at the household level is much more important. The results in column (3) ignore the replication of observations at the household level, whereas those in column (4) take it into account. (Both (3) and (4) ignore first-stage clustering and stratification.) For the German data, this leads to standard errors that are about twice as large. This shows that survey design *can* matter. However, the precise effect appears to depend on the survey analysed: corresponding estimates in columns (3) and (4) differ little when BHPS data are used, by contrast with the GSOEP case.

— Table 1 near here —

The fact that ignoring stratification and first-stage clustering has only a small impact might be interpreted as good news for practitioners using these data, or for those using surveys in which primary sampling unit and strata identification variables are not made available, but it is not clear that this empirical finding can be generalized. Whatever the case, our variance estimators provide a straightforward means by which researchers can accommodate a range of design effects in their analysis.

⁵Stata programs to compute the estimators are available from the authors on request.

5 References

- Biewen, M., 2002, Bootstrap Inference for Inequality, Mobility and Poverty Measurement. *Journal of Econometrics* 108, 317 – 342.
- Binder, D.A., 1983, On the Variances of Asymptotically Normal Estimators from Complex Surveys, *International Statistical Review* 51, 293 – 300.
- Binder, D.A. and Z. Patak, 1994, Use of Estimating Functions for Estimation From Complex Surveys, *Journal of the American Statistical Association* 89, 1035 – 1043.
- Cochran, W.G., 1977, *Sampling Techniques*, 3d ed. (Wiley New York).
- Cowell, F.A. , 1989, Sampling Variance and Decomposable Inequality Measures, *Journal of Econometrics* 42, 27 – 41.
- Cowell, F.A. and Jenkins, S.P., 2003, Estimating Welfare Indices: Household Weights and Sample Design, in: Y. Amiel and J.A. Bishop, eds., *Inequality, Welfare and Poverty: Theory and Measurement*, Volume 9 (Elsevier Science, Amsterdam) 147 – 172, .
- Deaton, A., 1997, *The Analysis of Household Surveys: a Microeconomic Approach to Development Policy* (Johns Hopkins University Press, Baltimore).
- Schluter, C. and M. Trede, 2002, Statistical Inference for Inequality and Poverty Measures with Dependent Data, *International Economic Review* 43, 185 – 200.
- SOEP-Group, 2001, The German Socio-Economic Panel (GSOEP) after more than 15 years - Overview, in: Holst, E., D.R. Lillard and T.A. DiPrete, eds., Proceedings of the 2000 Fourth International Conference of German Socio-Economic Panel Study Users (GSOEP2000), *Vierteljahreshefte zur Wirtschaftsforschung (Quarterly Journal of Economic Research)* Vol. 70, 7 – 14.
- Taylor, M., ed., 2002, *British Household Panel Survey User Manual*, Institute for Social and Economic Research, Colchester.
- Thistle, P.D. , 1990, Large Sample Properties of Two Inequality Indices, *Econometrica* 58, 725 – 728.
- Woodruff, R.S., 1971, A Simple Method for Approximating the Variance of a Complicated Estimate, *Journal of the American Statistical Association* 66, 411 – 414.

6 Tables

Table 1. Income inequality¹ in West Germany (1984) and Britain (1991)

Index	(1) Survey ²		(2) I.i.d. ³		(3) Survey, i.i.d. ⁴		(4) Survey ⁵	
	estimate	std. err.	estimate	std. err.	estimate	std. err.	estimate	std. err.
<i>German Socio-Economic Panel⁶ (1984)</i>								
GE(-1)	0.4397	0.2971	0.3992	0.2573	0.4397	0.1333	0.4397	0.2978
MLD	0.1339	0.0153	0.1338	0.0136	0.1339	0.0087	0.1339	0.0152
Theil	0.1540	0.0276	0.1540	0.0246	0.1540	0.0187	0.1540	0.0276
GE(2)	0.3673	0.1586	0.3502	0.1400	0.3673	0.1102	0.3673	0.1584
Atkinson(0.5)	0.0658	0.0079	0.0661	0.0071	0.0658	0.0052	0.0658	0.0079
Atkinson(1)	0.1253	0.0134	0.1253	0.0119	0.1253	0.0076	0.1253	0.0133
Atkinson(1.5)	0.2153	0.0477	0.2101	0.0419	0.2153	0.0219	0.2153	0.0478
Atkinson(2)	0.4679	0.1682	0.4439	0.1591	0.4679	0.0754	0.4679	0.1686
<i>British Household Panel Survey⁷ (1991)</i>								
GE(-1)	0.3656	0.0605	0.3627	0.0541	0.3656	0.0571	0.3655	0.0601
MLD	0.1907	0.0051	0.1914	0.0050	0.1907	0.0030	0.1906	0.0050
Theil	0.1779	0.0050	0.1784	0.0050	0.1779	0.0030	0.1779	0.0049
GE(2)	0.2064	0.0085	0.2071	0.0086	0.2064	0.0054	0.2064	0.0085
Atkinson(0.5)	0.0873	0.0022	0.0875	0.0022	0.0873	0.0013	0.0873	0.0021
Atkinson(1)	0.1736	0.0042	0.1741	0.0041	0.1736	0.0025	0.1736	0.0041
Atkinson(1.5)	0.2685	0.0084	0.2693	0.0082	0.2685	0.0058	0.2684	0.0082
Atkinson(2)	0.4223	0.0404	0.4204	0.0363	0.4223	0.0381	0.4223	0.0401

¹ Income refers to monthly equivalent household net income distributed among individuals (equivalence scale = square root of household size).

² Survey estimator, individual data, weight = individual sample weight, accounting for clustering and stratification; replication of observations at household level automatically accounted for.

³ I.i.d. estimator, household data, weight = household size * household sample weight (thus accounting for replication of observations at household level), ignoring clustering and stratification.

⁴ Survey estimator, individual data, weight = individual sample weight, ignoring clustering, stratification and replication of observations at household level (identical to i.i.d. estimator, individual data, weight = individual sample weight).

⁵ Survey estimator, individual data, weight = individual sample weight, ignoring clustering and stratification but accounting for replication of observations at household level (households are interpreted as clusters).

⁶ 110 strata, 516 clusters, 4232 households, 9441 individuals.

⁷ 75 strata, 250 clusters, 4814 households, 11616 individuals.