The Robustification of the Quintile Share Ratio

Beat Hulliger & Tobias Schoch

University of Applied Sciences Northwestern Switzerland (FHNW)

NTTS 2009 Conference
18th-20th February 2009, Brussels

SEVENTH FRAMEWORK PROGRAMME

Contents

1. Introduction
2. Methods
3. First Results
4. Outlook

Introduction

Laeken indicators

- European statistical **indicators on poverty and social exclusion**
- Lisbon Strategy: coordination of European social policies

Focus

Here only **monetary indicators**

Data

European Statistics on Income and Living Condition (EU-SILC), coordinated national surveys in all European countries including Switzerland.

AMELI

AMELI (Advanced Methodology for European Laeken Indicators) project on statistical problems including robustness, small area estimation and variance estimation. Development and evaluation by simulation. 2008-2011 FP7 SSH programme.

Motivation

Fact

Extreme incomes have potentially large influence on social indicators!

- The problems arising from the non-robustness are
  - biased estimates
  - inaccurate, misleading inference
- Robustness properties of welfare measures
  - Inequality measures (e.g. Gini coefficient) are known to be sensitive to outlying (extreme) observations.
  - Poverty measures are less influenced by extreme observations

NOT considered

Effect of admittely negative incomes

Quintile Share Ratio (QSR)

Description

QSR is defined as the ratio of total equivalized disposable income received by the 20% of a country's population with hightest income to that received by the 20% of the country's population with the lowest income.

Definition

$$\eta = \frac{\sum_{i \in U} y_i \mathbb{1}\{Q_{F_U}(0.8) < y_i \leq Q_{F_U}(1)\}}{\sum_{i \in U} y_i \mathbb{1}\{Q_{F_U}(0) \leq y_i \leq Q_{F_U}(0.2)\}} \qquad (1)$$

where $\mathbb{1}\{A\}$ is an indicator function for set A, F_U is the cdf, and $Q_{F_U}(\alpha) := F_U^{-1}(\alpha)$ the α quantile.

Quintile Share Ratio: Estimation

Definition

Let C_{F_S} be the (weighted) cumulative income functional

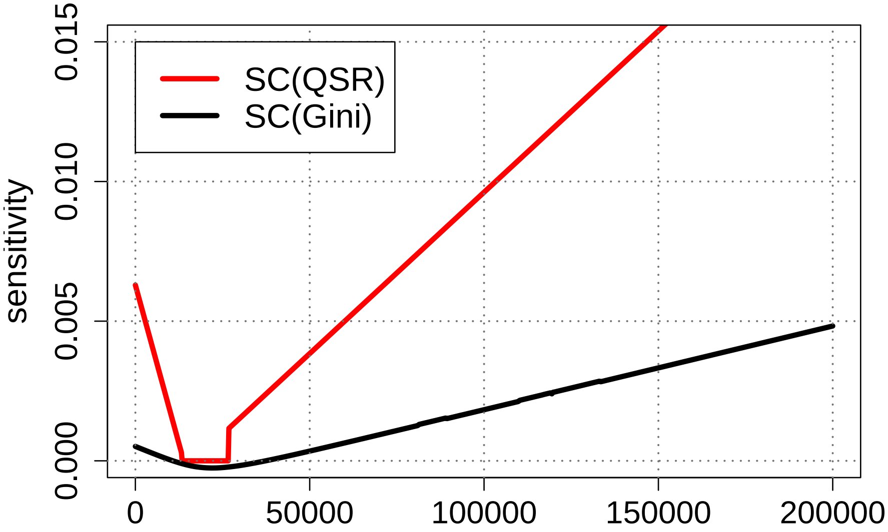
$$C_{F_S}(q) = \frac{\sum_{i \in S} w_i y_i \mathbb{1}\{y_i \leq Q_{F_S}(q)\}}{\sum_{i \in S} w_i \mathbb{1}\{y_i \leq Q_{F_S}(q)\}}, \tag{2}$$

where $0 \leq q \leq 1$, F_U is estimated by $F_S(t) = \sum_{i=1}^n w_i \mathbb{1}\{y_i \leq t\} / \sum_{i=1}^n w_i$ and by definition $C_{F_\mathbf{s}}(0) = 0$ and $C_{F_\mathbf{s}}(1) = \mu(y)$; (cf. Cowell and Victoria-Feser, 2002).

Hence, the QSR is be estimated as follows:

1. estimate $Q_{F_S}(0.2)$ and $Q_{F_S}(0.2)$
2. estimate the cumulative income functionals $m_1 = C_{F_s}(0.2)$ and $m_5 = C_{F_s}(1) - C_{F_s}(0.8)$
3. QSR is the ratio of $\hat{\eta} = \frac{m_5}{m_1} = \frac{C_{F_S}(1) - C_{F_S}(0.8)}{C_{F_S}(0.2)}$

Sensitivity Curve (Data: public use sample AT-SILC 2004)



Proposition 1: Trimmed Quintile Share Ratio (TQSR)

Definition TQSR

Extreme observations are (symmetrically) trimmed from both tails. Upper (α_u) and lower (α_l) trimming proportions are equal.

$$\hat{\eta}_{TQSR}(\alpha_l, \alpha_u) = \frac{C_{F_S}(1-\alpha_u) - C_{F_S}(0.8)}{C_{F_S}(0.2) - C_{F_S}(\alpha_l)}. \quad (3)$$

- Trimming: easy well-understood computation: direct intuitive appeal
- Cowell and Victoria-Feser (2006) discuss welfare analysis with trimmed data
- TQSR can lead to severe bias

Proposition 2: Bias-Corrected Quintile Share Ratio (BQSR1)

Definition BQSR1

$$\hat{\eta}_{BQSR1}(\alpha_l, \alpha_u) = \frac{C_{F_S}(1-\alpha_u) - C_{F_S}(0.8)}{C_{F_S}(0.2-\alpha_l)}$$

Note

Under a theoretical distribution α_l may be $\alpha_l = h(\alpha_u)$ such that

$$\hat{\eta}_{BQSE1}\left(h\left(\alpha_u\right), \alpha_u\right) = \hat{\eta}$$

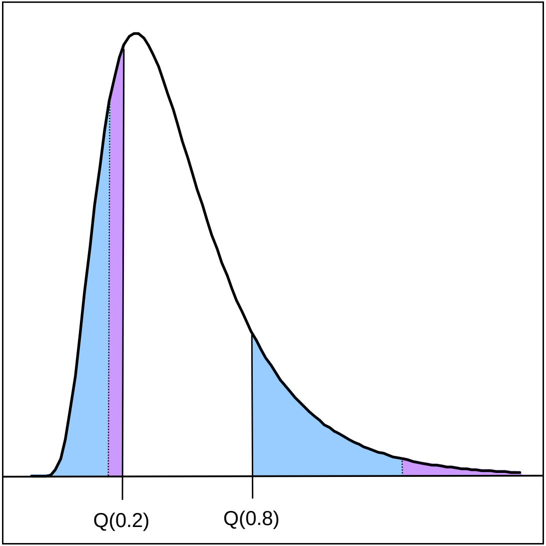


Figure: Idealized density function

Proposition 3: Bias-Corrected Quintile Share Ratio (BQSR2)

The idea:

- trim only m_5
- to compensate the bias trim twice
- and then "correct" accordingly

Definition BQSR2

$$\hat{\eta}_{BQSR2}(\alpha_u) = \frac{C_{F_S}(1-\alpha_u) - C_{F_S}(0.8)}{C_{F_S}(0.2)} \cdot \left[\frac{C_{F_S}(1-\alpha_u) - C_{F_S}(0.8)}{C_{F_S}(1-2\alpha_u) - C_{F_S}(0.8)} \right]$$

Proposition 4: M-Estimation Quintile Share Ratio (MQSR)

Definition MQSR

Use M-estimation with asymmetric (Huber) ψ-function $\psi(x,k) = min(x,k)$ with $k > 0$ to estimate m_5 the weighted mean of incomes above $Q_{F_S}(0.8)$. $T(F_S, k)$ is the solution of the estimating equation

$$\sum_{i \in S} w_i \, \mathbb{1}\{Q_{F_S}(0.8) < y_i\} \, \psi(y_i - T, k) = 0. \quad (4)$$

Choose k either by:

1. using a preliminary scale and judgment
2. minimizing an estimated MSE (i.e. MER-Estimator $T\left(F_S, k_0(F_S)\right)$)

Sensitivity Curves

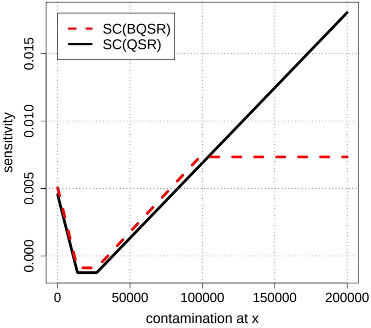


Figure: SC for BQSR1 vs. QSR

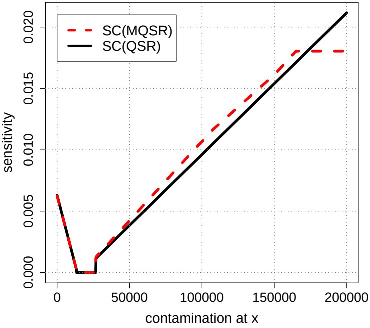


Figure: SC for MQSR vs. QSR

Preliminary Results 1

Table: Relative Bias (in % QSR) of TQSR and BQSR1

	TQSR (α_u)				BQSR1 (α_u)			
(α_l)$\bigtriangledown$	0.001	0.002	0.005	0.01	0.001	0.002	0.005	0.01
0	-1.72	-3.18	-5.31	-7.56	-1.72	-3.18	-5.31	-7.56
0.005	-4.02	-5.44	-7.52	-9.72	-0.65	-2.12	-4.28	-6.56
0.010	-5.95	-7.35	-9.39	-11.55	0.48	-1.01	-3.19	-5.49
0.020	-9.15	-10.49	-12.46	-14.55	2.73	1.21	-1.02	-3.37
0.050	-15.35	-16.60	-18.44	-20.38	11.05	9.40	6.99	4.45
0.100	-21.96	-23.12	-24.81	-26.60	32.92	30.96	28.07	25.02

Data: public use sample AT-SILC 2004

Preliminary Results 2

Table: Relative Bias (in % QSR) of BQSR2, TQSR and MQSR

	trimming proportion (α_u)							
	0	0.001	0.002	0.003	0.005	0.01	0.02	0.05
TQSR	0	-1.72	-3.18	-4.28	-5.31	-7.56	-10.28	-15.45
BQSR2	0	-0.24	-1.34	-2.63	-3.03	-4.76	-6.39	-10.01
MQSR	0	-1.22	-2.09	-2.57	-3.60	-4.85	-6.36	-7.57
MER-QSR	0	-1.21	–	–	–	–	–	–

Note: for MQSR *k* was chosen such that the number of declared outliers is equivalent to the trimming proportion of the other estimators. Data: public use sample AT-SILC 2004.

Outlook

1. QSR parameters are taken as given
2. Robustification $\Rightarrow$ bias due to asymmetry
3. Search for a good robustifiation $\Rightarrow$ propositions look promising
4. First results
   - Bounded influence function
   - Balancing the bias
5. Outlook
   - Evaluation in AMELI project by extensive simulation
   - Variance and efficiency?