Measuring Inequality with Ordinal data

Frank Cowell

http://darp.lse.ac.uk/cowell.htm

Università di Verona: Alba di Canazei
Winter School

January 2015

Outline

Outline

Introduction

- Ordinal data issue widespread in inequality analysis
- Many applications proceed just as though cardinal:
  - life satisfaction / inequality of happiness: Oswald and Wu (2011), Stevenson and Wolfers (2008b), Yang (2008)
  - health status: Van Doorslaer and Jones (2003)

Introduction

- Ordinal data issue widespread in inequality analysis
- Many applications proceed just as though cardinal:
  - life satisfaction / inequality of happiness: Oswald and Wu (2011), Stevenson and Wolfers (2008b), Yang (2008)
  - health status: Van Doorslaer and Jones (2003)
- Small literature that takes ordinal problem seriously
  - early approaches using 1st order dominance, the median
  - Abul Naga and Yalcin (2008,2010), Allison and Foster (2004), Zheng (2011)
  - but these have limitations

Introduction

- Ordinal data issue widespread in inequality analysis
- Many applications proceed just as though cardinal:
  - life satisfaction / inequality of happiness: Oswald and Wu (2011), Stevenson and Wolfers (2008b), Yang (2008)
  - health status: Van Doorslaer and Jones (2003)
- Small literature that takes ordinal problem seriously
  - early approaches using 1st order dominance, the median
  - Abul Naga and Yalcin (2008,2010), Allison and Foster (2004), Zheng (2011)
  - but these have limitations
- Present approach based on Cowell and Flachaire (2014)

Outline

Income Inequality

- 3 ingredients:
  - **"income":** family income, earnings, wealth $x \in X \subseteq \mathbb{R}$.
  - **"income-receiving unit":** n persons
  - **method of aggregation:** function $X^n \to \mathbb{R}$

Income Inequality

- 3 ingredients:
  - **“income”:** family income, earnings, wealth $x \in X \subseteq \mathbb{R}$.
  - **“income-receiving unit”:** n persons
  - **method of aggregation:** function $X^n \to \mathbb{R}$
- Usually work with $X^n_\mu \subset \mathbb{R}$

Income Inequality

- 3 ingredients:
  - **"income":** family income, earnings, wealth $x \in X \subseteq \mathbb{R}$.
  - **"income-receiving unit":** n persons
  - **method of aggregation:** function $X^n \to \mathbb{R}$
- Usually work with $X^n_\mu \subset \mathbb{R}$
- X^n_μ: Distributions obtainable from a given total income $n\mu$ using lump-sum transfers

Income Inequality

- 3 ingredients:
  - **“income”:** family income, earnings, wealth $x \in X \subseteq \mathbb{R}$.
  - **“income-receiving unit”:** n persons
  - **method of aggregation:** function $X^n \to \mathbb{R}$
- Usually work with $X^n_\mu \subset \mathbb{R}$
- X^n_μ: Distributions obtainable from a given total income $n\mu$ using lump-sum transfers
- Obviously can’t do that here: μ is undefined

Utility

Cardinalisation and inequality

- 3 ingredients:
  - **"income":** $u = U(x)$.
  - **"income-receiving unit":** n persons (as before)
  - **method of aggregation:** function $\mathbb{U}^n \to \mathbb{R}$

Utility

Cardinalisation and inequality

- 3 ingredients:
  - **"income":** $u = U(x)$.
  - **"income-receiving unit":** n persons (as before)
  - **method of aggregation:** function $\mathbb{U}^n \to \mathbb{R}$
- Problem of cardinalisation

Utility

Cardinalisation and inequality

- 3 ingredients:
  - **"income":** $u = U(x)$.
  - **"income-receiving unit":** n persons (as before)
  - **method of aggregation:** function $\mathbb{U}^n \to \mathbb{R}$
- Problem of cardinalisation
- But just assuming cardinal utility is no use
  - Already pointed out in Atkinson (1970)
  - Dalton (1920) suggested inequality of (cardinal) utility
  - But if, for all i, you multiply u_i by $\lambda \in (0,1)$ and add $\delta = \mu[1-\lambda]$...
  - ...this will automatically reduce measured inequality.

Utility

Cardinalisation and inequality

- 3 ingredients:
  - **"income":** $u = U(x)$.
  - **"income-receiving unit":** n persons (as before)
  - **method of aggregation:** function $\mathbb{U}^n \to \mathbb{R}$
- Problem of cardinalisation
- But just assuming cardinal utility is no use
  - Already pointed out in Atkinson (1970)
  - Dalton (1920) suggested inequality of (cardinal) utility
  - But if, for all i, you multiply u_i by $\lambda \in (0,1)$ and add $\delta = \mu[1-\lambda]$...
  - ...this will automatically reduce measured inequality.
- Is this just a technicality?
- Can we proceed just as with regular income?

Outline

Categorical variable

Example: Access to Services

	Case 1	Case 2
	n_k	n_k
Both Gas and Electricity	25	0
Electricity only	25	50
Gas only	25	50
Neither	25	0

Categorical variable

Example: Access to Services

	Case 1	Case 2
	n_k	n_k
Both Gas and Electricity	25	0
Electricity only	25	50
Gas only	25	50
Neither	25	0

- Suppose we have no information about needs / usage

Categorical variable

Example: Access to Services

	Case 1	Case 2
	n_k	n_k
Both Gas and Electricity	25	0
Electricity only	25	50
Gas only	25	50
Neither	25	0

- Suppose we have no information about needs / usage
- It seems clear that Case 1 is more unequal than Case 2

Example self-reported health

- World Health Survey (WHS)
    - a general population survey
    - developed by WHO

Example self-reported health

- World Health Survey (WHS)
  - a general population survey
  - developed by WHO
- Question: Health State Descriptions
  - overall health
  - including both physical and mental health

Example self-reported health

- World Health Survey (WHS)
  - a general population survey
  - developed by WHO
- Question: Health State Descriptions
  - overall health
  - including both physical and mental health
- In general, how would you rate your health today?
  - Very good
  - Good
  - Moderate
  - Bad
  - Very Bad

Example self-reported health

- World Health Survey (WHS)
  - a general population survey
  - developed by WHO
- Question: Health State Descriptions
  - overall health
  - including both physical and mental health
- In general, how would you rate your health today?
  - Very good
  - Good
  - Moderate
  - Bad
  - Very Bad
- Compare distributions across countries

SRH Results: four countries

	Austria	**UK**	**Mexico**	**Bangladesh**
	number of responses			
Very good	423	318	7193	494
Good	390	498	18112	1949
Moderate	200	278	11221	2132
Bad	36	82	2002	741
Very bad	4	17	218	228

SRH Results: four countries

	Austria	**UK**	**Mexico**	**Bangladesh**
	number of responses			
Very good	423	318	7193	494
Good	390	498	18112	1949
Moderate	200	278	11221	2132
Bad	36	82	2002	741
Very bad	4	17	218	228

- For all countries: rank categories in order

SRH Results: four countries

	Austria	**UK**	**Mexico**	**Bangladesh**
	number of responses			
Very good	423	318	7193	494
Good	390	498	18112	1949
Moderate	200	278	11221	2132
Bad	36	82	2002	741
Very bad	4	17	218	228

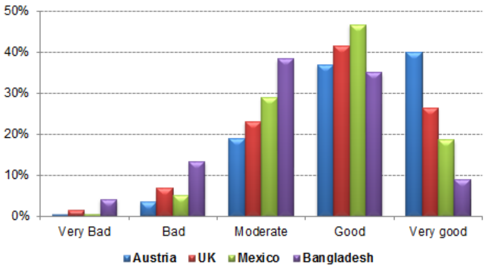
- For all countries: rank categories in order
- For each country: compute freq distributions across categories

SRH Results: four countries

	Austria	UK	Mexico	Bangladesh
	number of responses			
Very good	423	318	7193	494
Good	390	498	18112	1949
Moderate	200	278	11221	2132
Bad	36	82	2002	741
Very bad	4	17	218	228

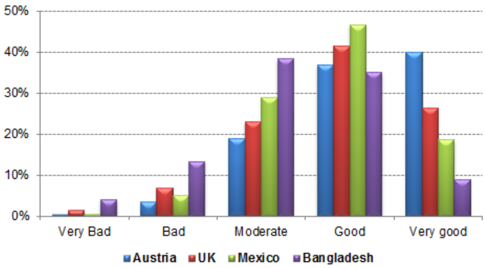
- For all countries: rank categories in order
- For each country: compute freq distributions across categories
- How to evaluate inequality?

SRH Inequality: Gini



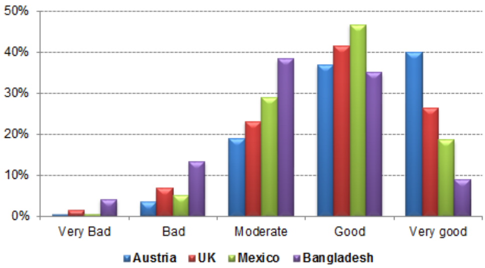
	At	UK	Mx	BD	
(1,2,3,4,5)	0.111	0.130	0.116	0.154	(BD,UK,Mx,At)

SRH Inequality: Gini



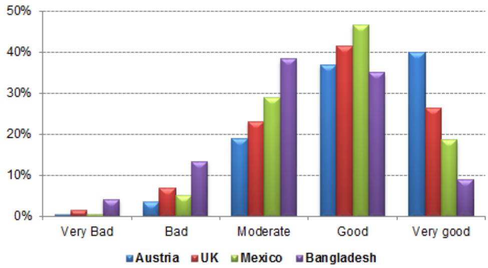
	At	UK	Mx	BD	
(1,2,3,4,5)	0.111	0.130	0.116	0.154	(BD,UK,Mx,At)
(1,2,3,4,1000)	0.593	0.725	0.800	0.884	(BD,Mx,UK,At)

SRH Inequality: Gini



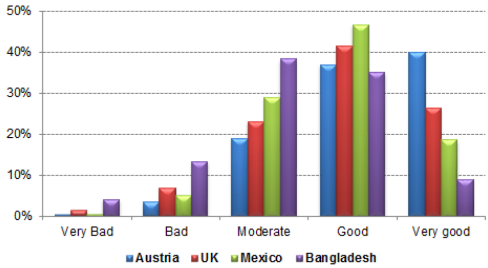
	At	UK	Mx	BD	
(1,2,3,4,5)	0.111	0.130	0.116	0.154	(BD,UK,Mx,At)
(1,2,3,4,1000)	0.593	0.725	0.800	0.884	(BD,Mx,UK,At)
(-1000,2,3,4,5)	0.608	0.821	0.856	2.377	(BD,Mx,UK,At)

SRH Inequality: Coeff of Variation



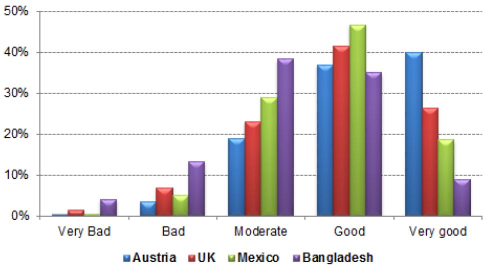
		At	UK	Mx	BD	
(1,2,3,4,5)		0.209	0.244	0.219	0.287	(BD,UK,Mx,At)

SRH Inequality: Coeff of Variation



	At	UK	Mx	BD	
(1,2,3,4,5)	0.209	0.244	0.219	0.287	(BD,UK,Mx,At)
(1,2,3,4,1000)	1.210	1.638	2.056	3.088	(BD,Mx,UK,At)

SRH Inequality: Coeff of Variation



	At	UK	Mx	BD	
(1,2,3,4,5)	0.209	0.244	0.219	0.287	(BD,UK,Mx,At)
(1,2,3,4,1000)	1.210	1.638	2.056	3.088	(BD,Mx,UK,At)
(-1000,2,3,4,5)	187.5	11.43	40.45	5.264	(At,Mx,UK,BD)

Outline

Status and Information

Status and Information

- Step 1 is to define status

Status and Information

- Step 1 is to define status
  - depends on the purpose of inequality analysis

Status and Information

- Step 1 is to define status
  - depends on the purpose of inequality analysis
  - depends on structure of information

Status and Information

- Step 1 is to define status
    - depends on the purpose of inequality analysis
    - depends on structure of information
    - conventional inequality approach only works in narrowly defined information structure

Status and Information

- Step 1 is to define status
  - depends on the purpose of inequality analysis
  - depends on structure of information
  - conventional inequality approach only works in narrowly defined information structure
- In some cases a person's status is self-defining
  - income
  - wealth

Status and Information

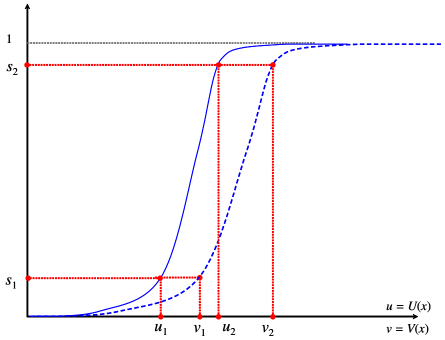
- Step 1 is to define status
  - depends on the purpose of inequality analysis
  - depends on structure of information
  - conventional inequality approach only works in narrowly defined information structure
- In some cases a person's status is self-defining
  - income
  - wealth
- In some cases defined given additional distribution-free information
  - example: if it is known that utility is $\log(x)$

Status and Information

- Step 1 is to define status
  - depends on the purpose of inequality analysis
  - depends on structure of information
  - conventional inequality approach only works in narrowly defined information structure
- In some cases a person's status is self-defining
  - income
  - wealth
- In some cases defined given additional distribution-free information
  - example: if it is known that utility is $\log(x)$
- In some cases requires information on distribution
  - GRE, TOEFL
  - "opportunity" (de Barros et al. 2008)

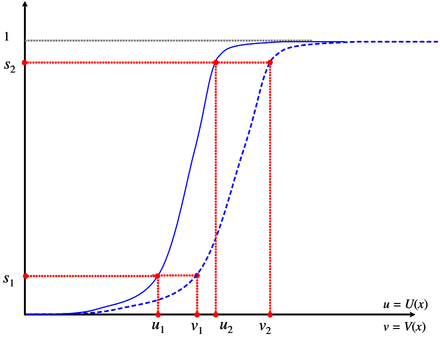
Status and Distribution (1)

- i's status uniquely defined for a given distribution of u



Status and Distribution (1)

- i's status uniquely defined for a given distribution of u



- disposes of the problem of cardinalisation
  - U and $V = \varphi(U)$ two cardinalisations of the utility of x
  - for each i:u_i and v_i map into s_i

Status and distribution (2)

Status and distribution (2)

- This approach works for categorical data
  - we just have an ordered arrangement of categories $1, 2, \ldots, k, \ldots, K$
  - and the numbers in each category $n_1, n_2, \ldots, n_k, \ldots, n_K$

Status and distribution (2)

- This approach works for categorical data
  - we just have an ordered arrangement of categories $1, 2, \ldots, k, \ldots, K$
  - and the numbers in each category $n_1, n_2, \ldots, n_k, \ldots, n_K$
- Merger principle
  - merge two adjacent categories that are irrelevant for i
  - then this should leave i's status unaltered

Status and distribution (2)

- This approach works for categorical data
  - we just have an ordered arrangement of categories $1, 2, ..., k, ..., K$
  - and the numbers in each category $n_1, n_2, ..., n_k, ..., n_K$
- Merger principle
  - merge two adjacent categories that are irrelevant for i
  - then this should leave i's status unaltered
- Principle implies that status should be additive in the n_k
  - downward-looking status: $\sum_{\ell=1}^{k(i)} n_\ell$
  - upward-looking status: $\sum_{\ell=k(i)}^{K} n_\ell$
  - see also Yitzhaki (1979)

Elements of the Model

- Individual's status is given by $s \in S \subseteq \mathbb{R}$
  - status determined from utility?

Elements of the Model

- Individual's status is given by $s \in S \subseteq \mathbb{R}$
  - status determined from utility?
- Vector of status in a population of size n : $\mathbf{s} \in S^n$

Elements of the Model

- Individual's status is given by $s \in S \subseteq \mathbb{R}$
  - status determined from utility?
- Vector of status in a population of size n : $\mathbf{s} \in S^n$
- $e \in S$: an equality-reference point

Elements of the Model

- Individual's status is given by $s \in S \subseteq \mathbb{R}$
  - status determined from utility?
- Vector of status in a population of size n : $\mathbf{s} \in S^n$
- $e \in S$: an equality-reference point
  - could be specified exogenously
  - could also depend on status vector $e = \eta(\mathbf{s})$
  - η need not be increasing in each component of $\mathbf{s}$

Elements of the Model

- Individual's status is given by $s \in S \subseteq \mathbb{R}$
  - status determined from utility?
- Vector of status in a population of size n : $\mathbf{s} \in S^n$
- $e \in S$: an equality-reference point
  - could be specified exogenously
  - could also depend on status vector $e = \eta(\mathbf{s})$
  - η need not be increasing in each component of $\mathbf{s}$
- Inequality: aggregate distance from e
  - don't need an explicit distance function
  - implicitly define through inequality ordering $\succeq$

Outline

Basic Axioms

- **[Continuity]** $\succeq$ is continuous on S^{n+1}

Basic Axioms

- **[Continuity]** $\succeq$ is continuous on S^{n+1}
- **[Monotonicity]** If $\mathbf{s}, \mathbf{s}' \in S^n$ differ only in their ith component then (a) if $s'_i \geq e$:$s_i > s'_i \iff (\mathbf{s}, e) \succ (\mathbf{s}', e)$; (b) if $s'_i \leq e$: $\iff (\mathbf{s}, e) \succ (\mathbf{s}', e)$

Basic Axioms

- **[Continuity]** $\succeq$ is continuous on S^{n+1}
- **[Monotonicity]** If $\mathbf{s}, \mathbf{s}' \in S^n$ differ only in their ith component then (a) if $s_i' \geq e$:$s_i > s_i' \iff (\mathbf{s}, e) \succ (\mathbf{s}', e)$; (b) if $s_i' \leq e$: $\iff (\mathbf{s}, e) \succ (\mathbf{s}', e)$
- **[Independence]** If $\mathbf{s}(\varsigma, i), \mathbf{s}'(\varsigma, i) \in S^n$ satisfy$(\mathbf{s}(\varsigma, i), e) \sim (\mathbf{s}'(\varsigma, i), e)$ for some ς then $(\mathbf{s}(\varsigma, i), e) \sim (\mathbf{s}'(\varsigma, i), e)$ for all ς

Basic Axioms

- **[Continuity]** $\succeq$ is continuous on S^{n+1}
- **[Monotonicity]** If $\mathbf{s}, \mathbf{s}' \in S^n$ differ only in their ith component then (a) if $s_i' \geq e : s_i > s_i' \iff (\mathbf{s}, e) \succ (\mathbf{s}', e)$; (b) if $s_i' \leq e$: $\iff (\mathbf{s}, e) \succ (\mathbf{s}', e)$
- **[Independence]** If $\mathbf{s}(\varsigma, i), \mathbf{s}'(\varsigma, i) \in S^n$ satisfy$(\mathbf{s}(\varsigma, i), e) \sim (\mathbf{s}'(\varsigma, i), e)$ for some ς then $(\mathbf{s}(\varsigma, i), e) \sim (\mathbf{s}'(\varsigma, i), e)$ for all ς
- **[Anonymity]** For all $\mathbf{s} \in S^n$ and permutation matrix Π: $(\Pi \mathbf{s}, e) \sim (\mathbf{s}, e)$

Standard result

Theorem

Continuity, Monotonicity, Independence, Anonymity jointly imply $\succeq$ *is representable by the continuous function* $I : S_e^n \to \mathbb{R}$ *where* $I(\mathbf{s}; e) = \Phi\left(\sum_{i=1}^{n} d(s_i, e), e\right)$, *where* $d : S \to \mathbb{R}$ *is a continuous function that is strictly increasing (decreasing) in its first argument if* $s_i > e$ $(s_i < e)$.

Standard result

Theorem

Continuity, Monotonicity, Independence, Anonymity jointly imply $\succeq$ *is representable by the continuous function* $I : S_e^n \rightarrow \mathbb{R}$ *where* $I(\mathbf{s}; e) = \Phi\left(\sum_{i=1}^{n} d(s_i, e), e\right)$, *where* $d : S \rightarrow \mathbb{R}$ *is a continuous function that is strictly increasing (decreasing) in its first argument if* $s_i > e$ $(s_i < e)$.

Corollary

Inequality is total "distance" from equality. Distance d is continuous. $d(s, e)$ *is increasing in status if you move away from the reference point.*

Structure Theorem

- We need more structure on the problem

Structure Theorem

- We need more structure on the problem
- **[Scale invariance 1]** For all $\lambda \in \mathbb{R}_+$: if $\mathbf{s}, \mathbf{s}', \lambda\mathbf{s}, \lambda\mathbf{s}' \in S^n$ and $e, e' \in S$ then $(\mathbf{s}, e) \sim (\mathbf{s}', e') \Rightarrow (\lambda\mathbf{s}, e) \sim (\lambda\mathbf{s}', e')$.
- **[Scale invariance 2]** For all $\lambda \in \mathbb{R}_+$: if $\mathbf{s}, \mathbf{s}', \lambda\mathbf{s}, \lambda\mathbf{s}' \in S^n$ and $e, e', \lambda e, \lambda e' \in S$ then $(\mathbf{s}, e) \sim (\mathbf{s}', e') \Rightarrow (\lambda\mathbf{s}, \lambda e) \sim (\lambda\mathbf{s}', \lambda e')$

Structure Theorem

- We need more structure on the problem
- **[Scale invariance 1]** For all $\lambda \in \mathbb{R}_+$: if $\mathbf{s}, \mathbf{s}', \lambda\mathbf{s}, \lambda\mathbf{s}' \in S^n$ and $e, e' \in S$ then $(\mathbf{s}, e) \sim (\mathbf{s}', e') \Rightarrow (\lambda\mathbf{s}, e) \sim (\lambda\mathbf{s}', e')$.
- **[Scale invariance 2]** For all $\lambda \in \mathbb{R}_+$: if $\mathbf{s}, \mathbf{s}', \lambda\mathbf{s}, \lambda\mathbf{s}' \in S^n$ and $e, e', \lambda e, \lambda e' \in S$ then $(\mathbf{s}, e) \sim (\mathbf{s}', e') \Rightarrow (\lambda\mathbf{s}, \lambda e) \sim (\lambda\mathbf{s}', \lambda e')$

Theorem

Impose also Scale irrelevance 1. Then $d(s, e) = A(e)\, s^{\alpha(e)}$

Structure Theorem

- We need more structure on the problem
- **[Scale invariance 1]** For all $\lambda \in \mathbb{R}_+$: if $\mathbf{s}, \mathbf{s}', \lambda\mathbf{s}, \lambda\mathbf{s}' \in S^n$ and $e, e' \in S$ then $(\mathbf{s}, e) \sim (\mathbf{s}', e') \Rightarrow (\lambda\mathbf{s}, e) \sim (\lambda\mathbf{s}', e')$.
- **[Scale invariance 2]** For all $\lambda \in \mathbb{R}_+$: if $\mathbf{s}, \mathbf{s}', \lambda\mathbf{s}, \lambda\mathbf{s}' \in S^n$ and $e, e', \lambda e, \lambda e' \in S$ then $(\mathbf{s}, e) \sim (\mathbf{s}', e') \Rightarrow (\lambda\mathbf{s}, \lambda e) \sim (\lambda\mathbf{s}', \lambda e')$

Theorem

Impose also Scale irrelevance 1. Then $d(s,e) = A(e)\, s^{\alpha(e)}$

Theorem

Impose instead Scale Invariance 2. Then $d(s,e) = e^{\beta} \phi\left(\frac{s}{e}\right)$ *.where* β *is a constant and* ϕ *is arbitrary*

Structure Theorem

- We need more structure on the problem
- **[Scale invariance 1]** For all $\lambda \in \mathbb{R}_+$: if $\mathbf{s}, \mathbf{s}', \lambda\mathbf{s}, \lambda\mathbf{s}' \in S^n$ and $e, e' \in S$ then $(\mathbf{s}, e) \sim (\mathbf{s}', e') \Rightarrow (\lambda\mathbf{s}, e) \sim (\lambda\mathbf{s}', e')$.
- **[Scale invariance 2]** For all $\lambda \in \mathbb{R}_+$: if $\mathbf{s}, \mathbf{s}', \lambda\mathbf{s}, \lambda\mathbf{s}' \in S^n$ and $e, e', \lambda e, \lambda e' \in S$ then $(\mathbf{s}, e) \sim (\mathbf{s}', e') \Rightarrow (\lambda\mathbf{s}, \lambda e) \sim (\lambda\mathbf{s}', \lambda e')$

Theorem

Impose also Scale irrelevance 1. Then $d(s,e) = A(e)\, s^{\alpha(e)}$

Theorem

Impose instead Scale Invariance 2. Then $d(s,e) = e^{\beta}\phi\left(\frac{s}{e}\right)$ *.where* β *is a constant and* ϕ *is arbitrary*

Corollary

Inequality represented as $I_\alpha(\mathbf{s}; e) := \frac{1}{\alpha[\alpha-1]}\left[\frac{1}{n}\sum_{i=1}^{n} s_i^{\alpha} - e^{\alpha}\right]$

Outline

A usable inequality index?

- A *class* of functions available as inequality measures:
  - $\Phi(I_\alpha(\mathbf{s};e),e)$
  - $e=\eta(\mathbf{s})$, the reference point
  - $I_\alpha(\mathbf{s};e) := \frac{1}{\alpha[\alpha-1]}\left[\frac{1}{n}\sum_{i=1}^{n} s_i^\alpha - e^\alpha\right]$

A usable inequality index?

- A *class* of functions available as inequality measures:
  - $\Phi(I_\alpha(\mathbf{s};e), e)$
  - $e = \eta(\mathbf{s})$, the reference point
  - $I_\alpha(\mathbf{s};e) := \frac{1}{\alpha[\alpha-1]}\left[\frac{1}{n}\sum_{i=1}^{n} s_i^\alpha - e^\alpha\right]$
- Do functions $\Phi(I_\alpha(\mathbf{s};e), e)$ "look like" inequality measures?
  - transfer principle?
  - reference point?
  - sensitivity to parameters
- What is the appropriate form for Φ?
  - may depend on the reference status e
  - may depend on interpretation

Outline

Four distributional scenarios (1)

	Case 0		Case 1		Case 2		Case 3	
	n_k	s_i	n_k	s_i	n_k	s_i	n_k	s_i
B	0		25	1	0		25	1
E	50	1	25	$3/4$	50	1	25	$3/4$
G	25	$1/2$	25	$1/2$	50	$1/2$	50	$1/2$
N	25	$1/4$	25	$1/4$	0		0	
$\mu(\mathbf{s})$		$11/16$		$5/8$		$3/4$		$11/16$

Four distributional scenarios (1)

	Case 0		Case 1		Case 2		Case 3	
	n_k	s_i	n_k	s_i	n_k	s_i	n_k	s_i
B	0		25	1	0		25	1
E	50	1	25	$3/4$	50	1	25	$3/4$
G	25	$1/2$	25	$1/2$	50	$1/2$	50	$1/2$
N	25	$1/4$	25	$1/4$	0		0	
$\mu(\mathbf{s})$		$11/16$		$5/8$		$3/4$		$11/16$

- n_k is # persons in category $k \in \{B, E, G, N\}$

Four distributional scenarios (1)

	Case 0		Case 1		Case 2		Case 3	
	n_k	s_i	n_k	s_i	n_k	s_i	n_k	s_i
B	0		25	1	0		25	1
E	50	1	25	3/4	50	1	25	3/4
G	25	1/2	25	1/2	50	1/2	50	1/2
N	25	1/4	25	1/4	0		0	
$\mu(\mathbf{s})$		11/16		5/8		3/4		11/16

- n_k is # persons in category $k \in \{B, E, G, N\}$
- $s_i = \frac{1}{n} \sum_{\ell=1}^{k(i)} n_\ell$ – *downward*-looking status

Four distributional scenarios (1′)

	Case 0		Case 1		Case 2		Case 3	
	n_k	s'_i	n_k	s'_i	n_k	s'_i	n_k	s'_i
B	0		25	$1/4$	0		25	$1/4$
E	50	$1/2$	25	$1/2$	50	$1/2$	25	$1/2$
G	25	$3/4$	25	$3/4$	50	1	50	1
N	25	1	25	1	0		0	
$\mu(\mathbf{s})$		$11/16$		$5/8$		$3/4$		$11/16$

Four distributional scenarios (1′)

	Case 0		Case 1		Case 2		Case 3	
	n_k	s'_i	n_k	s'_i	n_k	s'_i	n_k	s'_i
B	0		25	1/4	0		25	1/4
E	50	1/2	25	1/2	50	1/2	25	1/2
G	25	3/4	25	3/4	50	1	50	1
N	25	1	25	1	0		0	
$\mu(\mathbf{s})$		11/16		5/8		3/4		11/16

- n_k is # persons in category $k \in \{B, E, G, N\}$

Four distributional scenarios (1′)

	Case 0		Case 1		Case 2		Case 3	
	n_k	s'_i	n_k	s'_i	n_k	s'_i	n_k	s'_i
B	0		25	$1/4$	0		25	$1/4$
E	50	$1/2$	25	$1/2$	50	$1/2$	25	$1/2$
G	25	$3/4$	25	$3/4$	50	1	50	1
N	25	1	25	1	0		0	
$\mu(\mathbf{s})$		$11/16$		$5/8$		$3/4$		$11/16$

- n_k is # persons in category $k \in \{\mathrm{B}, \mathrm{E}, \mathrm{G}, \mathrm{N}\}$
- $s'_i = \frac{1}{n} \sum_{\ell=k(i)}^{K} n_\ell$ – *upward*-looking status

Four distributional scenarios (2)

	Case 0		Case 1		Case 2		Case 3	
	n_k	s_i	n_k	s_i	n_k	s_i	n_k	s_i
B	0		25	1	0		25	1
E	50	1	25	$3/4$	50	1	25	$3/4$
G	25	$1/2$	25	$1/2$	50	$1/2$	50	$1/2$
N	25	$1/4$	25	$1/4$	0		0	
$\mu(\mathbf{s})$		$11/16$		$5/8$		$3/4$		$11/16$

- Case 0 to Case 1:

Four distributional scenarios (2)

	Case 0		Case 1		Case 2		Case 3	
	n_k	s_i	n_k	s_i	n_k	s_i	n_k	s_i
B	0		25	1	0		25	1
E	50	1	25	$3/4$	50	1	25	$3/4$
G	25	$1/2$	25	$1/2$	50	$1/2$	50	$1/2$
N	25	$1/4$	25	$1/4$	0		0	
$\mu(\mathbf{s})$		$11/16$		$5/8$		$3/4$		$11/16$

- Case 0 to Case 1:
  - 25 people promoted from E to B
  - if e equals to any of values taken by $\mu(\mathbf{s})$
  - then inequality increases

Four distributional scenarios (3)

	Case 0		Case 1		Case 2		Case 3	
	n_k	s_i	n_k	s_i	n_k	s_i	n_k	s_i
B	0		25	1	0		25	1
E	50	1	25	$3/4$	50	1	25	$3/4$
G	25	$1/2$	25	$1/2$	50	$1/2$	50	$1/2$
N	25	$1/4$	25	$1/4$	0		0	
$\mu(\mathbf{s})$		$11/16$		$5/8$		$3/4$		$11/16$

- Case 0 to Case 2:

Four distributional scenarios (3)

	Case 0		Case 1		Case 2		Case 3	
	n_k	s_i	n_k	s_i	n_k	s_i	n_k	s_i
B	0		25	1	0		25	1
E	50	1	25	$3/4$	50	1	25	$3/4$
G	25	$1/2$	25	$1/2$	50	$1/2$	50	$1/2$
N	25	$1/4$	25	$1/4$	0		0	
$\mu(\mathbf{s})$		$11/16$		$5/8$		$3/4$		$11/16$

- Case 0 to Case 2:
  - 25 people promoted from N to G
  - if e equals to any of values taken by $\mu(\mathbf{s})$
  - then inequality decreases

“Transfer Principle”?

"Transfer Principle"?

	Case 0		Case 1		Case 2		Case 3	
	n_k	s_i	n_k	s_i	n_k	s_i	n_k	s_i
B	0		25	1	0		25	1
E	50	1	25	$3/4$	50	1	25	$3/4$
G	25	$1/2$	25	$1/2$	50	$1/2$	50	$1/2$
N	25	$1/4$	25	$1/4$	0		0	
$\mu(\mathbf{s})$		$11/16$		$5/8$		$3/4$		$11/16$

"Transfer Principle"?

	Case 0		Case 1		Case 2		Case 3	
	n_k	s_i	n_k	s_i	n_k	s_i	n_k	s_i
B	0		25	1	0		25	1
E	50	1	25	$3/4$	50	1	25	$3/4$
G	25	$1/2$	25	$1/2$	50	$1/2$	50	$1/2$
N	25	$1/4$	25	$1/4$	0		0	
$\mu(\mathbf{s})$		$11/16$		$5/8$		$3/4$		$11/16$

- Case 0 to Case 1: inequality increases

"Transfer Principle"?

	Case 0		Case 1		Case 2		Case 3	
	n_k	s_i	n_k	s_i	n_k	s_i	n_k	s_i
B	0		25	1	0		25	1
E	50	1	25	$3/4$	50	1	25	$3/4$
G	25	$1/2$	25	$1/2$	50	$1/2$	50	$1/2$
N	25	$1/4$	25	$1/4$	0		0	
$\mu(\mathbf{s})$		$11/16$		$5/8$		$3/4$		$11/16$

- Case 0 to Case 1: inequality increases
- Case 0 to Case 2: inequality decreases

"Transfer Principle"?

	Case 0		Case 1		Case 2		Case 3	
	n_k	s_i	n_k	s_i	n_k	s_i	n_k	s_i
B	0		25	1	0		25	1
E	50	1	25	$^3/_4$	50	1	25	$^3/_4$
G	25	$^1/_2$	25	$^1/_2$	50	$^1/_2$	50	$^1/_2$
N	25	$^1/_4$	25	$^1/_4$	0		0	
$\mu(\mathbf{s})$		$^{11}/_{16}$		$^5/_8$		$^3/_4$		$^{11}/_{16}$

- Case 0 to Case 1: inequality increases
- Case 0 to Case 2: inequality decreases
- Case 0 to Case 3: combination results in ambiguous change

Outline

Reference point

Reference point

- **Mean status:** $e = \eta(\mathbf{s}) = \mu(\mathbf{s})$
  - for continuous distributions will equal 0.5
  - for categorical data, there is no counterpart to fixed-mean assumption in income-inequality analysis

Reference point

- **Mean status:** $e = \eta(\mathbf{s}) = \mu(\mathbf{s})$
  - for continuous distributions will equal 0.5
  - for categorical data, there is no counterpart to fixed-mean assumption in income-inequality analysis

- **Median status**: $e = \eta(\mathbf{s}) = \text{med}(\mathbf{s})$
  - not well-defined: any value in interval $M(\mathbf{s})$
  - $M(\mathbf{s}) = [1/2, 1)$ in cases 0 and 2
  - $M(\mathbf{s}) = [1/2, 3/4)$ in cases 1 and 3

Reference point

- **Mean status:** $e = \eta(\mathbf{s}) = \mu(\mathbf{s})$
  - for continuous distributions will equal 0.5
  - for categorical data, there is no counterpart to fixed-mean assumption in income-inequality analysis
- **Median status**: $e = \eta(\mathbf{s}) = \text{med}(\mathbf{s})$
  - not well-defined: any value in interval $M(\mathbf{s})$
  - $M(\mathbf{s}) = [1/2, 1)$ in cases 0 and 2
  - $M(\mathbf{s}) = [1/2, 3/4)$ in cases 1 and 3
- **Max status**: $e = 1$
  - for constant e this is only value that makes sense

Reference point

- **Mean status:** $e = \eta(\mathbf{s}) = \mu(\mathbf{s})$
  - for continuous distributions will equal 0.5
  - for categorical data, there is no counterpart to fixed-mean assumption in income-inequality analysis
- **Median status**: $e = \eta(\mathbf{s}) = \text{med}(\mathbf{s})$
  - not well-defined: any value in interval $M(\mathbf{s})$
  - $M(\mathbf{s}) = [1/2, 1)$ in cases 0 and 2
  - $M(\mathbf{s}) = [1/2, 3/4)$ in cases 1 and 3
- **Max status**: $e = 1$
  - for constant e this is only value that makes sense
- **Min status**: $e = 0$
  - counterpart for peer-exclusive case

Sensitivity

- α captures the sensitivity of measured inequality

Sensitivity

- α captures the sensitivity of measured inequality
- If α is high $I_\alpha(\mathbf{s};e) = \frac{1}{\alpha[\alpha-1]}\left[\frac{1}{n}\sum_{i=1}^{n} s_i^\alpha - e^\alpha\right]$, sensitive to high status-inequality

Sensitivity

- α captures the sensitivity of measured inequality
- If α is high $I_{\alpha}\left(\mathbf{s};e\right)=\frac{1}{\alpha[\alpha-1]}\left[\frac{1}{n}\sum_{i=1}^{n}s_i^{\alpha}-e^{\alpha}\right]$, sensitive to high status-inequality
- If $\alpha=0$ then $I_0\left(\mathbf{s};e\right)=-\frac{1}{n}\sum_{i=1}^{n}\log s_i+\log e$,

Sensitivity

- α captures the sensitivity of measured inequality
- If α is high $I_\alpha(\mathbf{s};e) = \frac{1}{\alpha[\alpha-1]}\left[\frac{1}{n}\sum_{i=1}^{n} s_i^\alpha - e^\alpha\right]$, sensitive to high status-inequality
- If $\alpha = 0$ then $I_0(\mathbf{s};e) = -\frac{1}{n}\sum_{i=1}^{n}\log s_i + \log e$,
- If $e = \mu(\mathbf{s})$ and $\alpha = 1$ then $\frac{1}{n}\sum_{i=1}^{n} s_i \log s_i - e\log e$

Behaviour of $I_0(\mathbf{s}; e)$

	Case 0	Case 1	Case 2	Case 3
$\mu(\mathbf{s})$	$11/16$	$5/8$	$3/4$	$11/16$
$\text{med}_1(\mathbf{s})$	$3/4$	$5/8$	$3/4$	$5/8$
$\text{med}_2(\mathbf{s})$	$1/2$	$1/2$	$1/2$	$1/2$
$I_0(\mathbf{s};\ \mu(\mathbf{s}))$	0.1451	0.1217	0.0588	0.0438
$I_0(\mathbf{s};\ \text{med}_1(\mathbf{s}))$	0.2321	0.1217	0.0588	-0.0515
$I_0(\mathbf{s};\ \text{med}_2(\mathbf{s}))$	-0.1732	-0.1013	-0.3465	-0.2746
$I_0(\mathbf{s};\ 1)$	0.5198	0.5917	0.3465	0.4184

Behaviour of $I_0(\mathbf{s}; e)$

	Case 0	Case 1	Case 2	Case 3
$\mu(\mathbf{s})$	$11/16$	$5/8$	$3/4$	$11/16$
$\text{med}_1(\mathbf{s})$	$3/4$	$5/8$	$3/4$	$5/8$
$\text{med}_2(\mathbf{s})$	$1/2$	$1/2$	$1/2$	$1/2$
$I_0(\mathbf{s};\ \mu(\mathbf{s}))$	0.1451	0.1217	0.0588	0.0438
$I_0(\mathbf{s};\ \text{med}_1(\mathbf{s}))$	0.2321	0.1217	0.0588	-0.0515
$I_0(\mathbf{s};\ \text{med}_2(\mathbf{s}))$	-0.1732	-0.1013	-0.3465	-0.2746
$I_0(\mathbf{s};\ 1)$	0.5198	0.5917	0.3465	0.4184

- $I_0(\mathbf{s};\ \mu(\mathbf{s}))$, $I_0(\mathbf{s};\ \text{med}_1(\mathbf{s}))$: inequality *decreases* for
  - Case 0 to 1, or Case 2 to 3
  - movement changes both the $\mu(\mathbf{s})$ and $\text{med}_1(\mathbf{s})$ ref points

Behaviour of $I_0(\mathbf{s}; e)$

	Case 0	Case 1	Case 2	Case 3
$\mu(\mathbf{s})$	$^{11}/_{16}$	$^5/_8$	$^3/_4$	$^{11}/_{16}$
$\text{med}_1(\mathbf{s})$	$^3/_4$	$^5/_8$	$^3/_4$	$^5/_8$
$\text{med}_2(\mathbf{s})$	$^1/_2$	$^1/_2$	$^1/_2$	$^1/_2$
$I_0(\mathbf{s};\ \mu(\mathbf{s}))$	0.1451	0.1217	0.0588	0.0438
$I_0(\mathbf{s};\ \text{med}_1(\mathbf{s}))$	0.2321	0.1217	0.0588	-0.0515
$I_0(\mathbf{s};\ \text{med}_2(\mathbf{s}))$	-0.1732	-0.1013	-0.3465	-0.2746
$I_0(\mathbf{s};\ 1)$	0.5198	0.5917	0.3465	0.4184

- $I_0(\mathbf{s};\ \mu(\mathbf{s}))$, $I_0(\mathbf{s};\ \text{med}_1(\mathbf{s}))$: inequality *decreases* for
  - Case 0 to 1, or Case 2 to 3
  - movement changes both the $\mu(\mathbf{s})$ and $\text{med}_1(\mathbf{s})$ ref points
- $I_0(\mathbf{s};\ \text{med}_2(\mathbf{s})) < 0$ for *all* cases in example!

Behaviour of $I_0\left(\mathbf{s}; e\right)$

	Case 0	Case 1	Case 2	Case 3
$\mu(\mathbf{s})$	$^{11}/_{16}$	$^{5}/_{8}$	$^{3}/_{4}$	$^{11}/_{16}$
$\mathrm{med}_1(\mathbf{s})$	$^{3}/_{4}$	$^{5}/_{8}$	$^{3}/_{4}$	$^{5}/_{8}$
$\mathrm{med}_2(\mathbf{s})$	$^{1}/_{2}$	$^{1}/_{2}$	$^{1}/_{2}$	$^{1}/_{2}$
$I_0(\mathbf{s};\ \mu(\mathbf{s}))$	0.1451	0.1217	0.0588	0.0438
$I_0(\mathbf{s};\ \mathrm{med}_1(\mathbf{s}))$	0.2321	0.1217	0.0588	-0.0515
$I_0(\mathbf{s};\ \mathrm{med}_2(\mathbf{s}))$	-0.1732	-0.1013	-0.3465	-0.2746
$I_0(\mathbf{s};\ 1)$	0.5198	0.5917	0.3465	0.4184

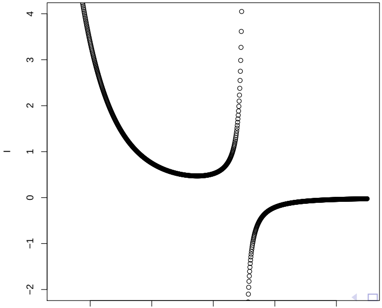
- $I_0(\mathbf{s};\ \mu(\mathbf{s}))$, $I_0(\mathbf{s};\ \mathrm{med}_1(\mathbf{s}))$: inequality *decreases* for
    - Case 0 to 1, or Case 2 to 3
    - movement changes both the $\mu(\mathbf{s})$ and $\mathrm{med}_1(\mathbf{s})$ ref points
- $I_0(\mathbf{s};\ \mathrm{med}_2(\mathbf{s})) < 0$ for *all* cases in example!
- But $I_0(\mathbf{s}; 1)$ seems sensible

Inequality measure

- For ordinal data, peer-inclusive status

Inequality measure

- For ordinal data, peer-inclusive status
- $I_\alpha(\mathbf{s},1) = \begin{cases} \frac{1}{\alpha(\alpha-1)}\left[\frac{1}{n}\sum_{i=1}^{n} s_i^\alpha - 1\right], & \text{if } \alpha \neq 0,\ \alpha<1 \\ -\frac{1}{n}\sum_{i=1}^{n}\log s_i. & \text{if } \alpha=0 \end{cases}$



Outline

Implementation

- Description of sample

$$x_i = \begin{cases} 1 & \text{with sample proportion } p_1 \\ 2 & \text{with sample proportion } p_2 \\ \dots & \\ K & \text{with sample proportion } p_K \end{cases},$$

Implementation

- Description of sample

$$x_i = \begin{cases} 1 & \text{with sample proportion } p_1 \\ 2 & \text{with sample proportion } p_2 \\ \dots & \\ K & \text{with sample proportion } p_K \end{cases},$$

- Point estimate of the index:
  - $I_\alpha = \begin{cases} \frac{1}{\alpha(\alpha-1)} \left[\sum_{i=1}^{K} p_i \left[\sum_{j=1}^{i} p_j \right]^\alpha - 1 \right] & \text{if } \alpha \neq 0,1 \\ -\sum_{i=1}^{K} p_i \log \left[\sum_{j=1}^{i} p_j \right] & \text{if } \alpha = 0 \end{cases}$
  - function of K parameter estimates $(p_1, p_2, \dots, p_K)$ following a multinomial

Asymptotics

Asymptotics

- From the CLT I_α is asymptotically Normally distributed

Asymptotics

- From the CLT I_α is asymptotically Normally distributed
- Estimator of cov matrix of $(p_1, p_2, \ldots, p_k)$ is

$$\Sigma = \frac{1}{n}\begin{bmatrix} p_1(1-p_1) & -p_1p_2 & \ldots & -p_1p_K \\ -p_2p_1 & p_2(1-p_2) & \ldots & -p_2p_K \\ \vdots & \vdots & \vdots & \vdots \\ -p_Kp_1 & -p_Kp_2 & \ldots & p_K(1-p_K) \end{bmatrix}$$

Asymptotics

- From the CLT I_α is asymptotically Normally distributed
- Estimator of cov matrix of $(p_1, p_2, \dots, p_k)$ is

$$\Sigma = \frac{1}{n}\begin{bmatrix} p_1(1-p_1) & -p_1p_2 & \dots & -p_1p_K \\ -p_2p_1 & p_2(1-p_2) & \dots & -p_2p_K \\ \vdots & \vdots & \vdots & \vdots \\ -p_Kp_1 & -p_Kp_2 & \dots & p_K(1-p_K) \end{bmatrix}$$

- $\widehat{\text{Var}}(I_\alpha) = D\Sigma D^\top \qquad \text{with} \qquad D = \left[\frac{\partial I_\alpha}{\partial p_1};\ \frac{\partial I_\alpha}{\partial p_2};\dots;\ \frac{\partial I_\alpha}{\partial p_K}\right]$
- $\frac{\partial I_\alpha}{\partial p_l} = \frac{1}{\alpha(\alpha-1)}\left(\left[\sum_{i=1}^{l} p_i\right]^\alpha + \alpha\sum_{i=l}^{K-1} p_i\left[\sum_{j=1}^{i} p_j\right]^{\alpha-1}\right), \alpha \neq 0$
- $\frac{\partial I_0}{\partial p_l} = -\log\left[\sum_{j=1}^{l} p_j\right] - \sum_{i=l}^{K-1} p_i\left[\sum_{j=1}^{i} p_j\right]^{-1}$

Outline

Confidence Intervals

Confidence Intervals

- 3 variants of CIs: Asymptotic, Percentile Bootstrap, Studentized Bootstrap
- $CI_{asym} = [I_\alpha - c_{0.975}\widehat{\text{Var}}(I_\alpha)^{1/2}; I_\alpha + c_{0.975}\widehat{\text{Var}}(I_\alpha)^{1/2}]$
  - $c_{0.975}$ from the Student distribution $T(n-1)$
  - do not always perform well in finite samples

Confidence Intervals

- 3 variants of CIs: Asymptotic, Percentile Bootstrap, Studentized Bootstrap
- $CI_{asym} = [I_\alpha - c_{0.975}\widehat{\text{Var}}(I_\alpha)^{1/2}; I_\alpha + c_{0.975}\widehat{\text{Var}}(I_\alpha)^{1/2}]$
  - $c_{0.975}$ from the Student distribution $T(n-1)$
  - do not always perform well in finite samples
- Bootstraps: generate resamples, $b = 1, \ldots, B$
  - for each resample b compute the inequality index
  - obtain B bootstrap statistics, I_α^b
  - also B bootstrap t-statistics $t_\alpha^b = (I_\alpha^b - I_\alpha)/\widehat{\text{Var}}(I_\alpha^b)^{1/2}$

Confidence Intervals

- 3 variants of CIs: Asymptotic, Percentile Bootstrap, Studentized Bootstrap
- $CI_{asym} = [I_\alpha - c_{0.975}\widehat{\text{Var}}(I_\alpha)^{1/2}\,;\ I_\alpha + c_{0.975}\widehat{\text{Var}}(I_\alpha)^{1/2}]$
  - $c_{0.975}$ from the Student distribution $T(n-1)$
  - do not always perform well in finite samples
- Bootstraps: generate resamples, $b = 1, \dots, B$
  - for each resample b compute the inequality index
  - obtain B bootstrap statistics, I_α^b
  - also B bootstrap t-statistics $t_\alpha^b = (I_\alpha^b - I_\alpha)/\widehat{\text{Var}}(I_\alpha^b)^{1/2}$
- $CI_{perc} = [c_{0.025}^b\,;\ c_{0.975}^b]$
  - $c_{0.025}^b$ and $c_{0.975}^b$ are from EDF of bootstrap statistics

Confidence Intervals

- 3 variants of CIs: Asymptotic, Percentile Bootstrap, Studentized Bootstrap
- $CI_{asym} = [I_\alpha - c_{0.975}\widehat{\mathrm{Var}}(I_\alpha)^{1/2}\,;\ I_\alpha + c_{0.975}\widehat{\mathrm{Var}}(I_\alpha)^{1/2}]$
  - $c_{0.975}$ from the Student distribution $T(n-1)$
  - do not always perform well in finite samples
- Bootstraps: generate resamples, $b = 1, \ldots, B$
  - for each resample b compute the inequality index
  - obtain B bootstrap statistics, I_α^b
  - also B bootstrap t-statistics $t_\alpha^b = (I_\alpha^b - I_\alpha)/\widehat{\mathrm{Var}}(I_\alpha^b)^{1/2}$
- $CI_{perc} = [c^b_{0.025}\,;\ c^b_{0.975}]$
  - $c^b_{0.025}$ and $c^b_{0.975}$ are from EDF of bootstrap statistics
- $CI_{stud} = [I_\alpha - c^*_{0.975}\widehat{\mathrm{Var}}(I_\alpha)^{1/2}\,;\ I_\alpha - c^*_{0.025}\widehat{\mathrm{Var}}(I_\alpha)^{1/2}]$
  - $c^*_{0.025}$ and $c^*_{0.975}$ are from EDF of the bootstrap t-statistics

Performance Test

Performance Test

- Take an example with 3 ordered categories ($K = 3$)

Performance Test

- Take an example with 3 ordered categories ($K = 3$)
- Samples are drawn from a multinomial distribution with probabilities $\pi = (0.3, 0.5, 0.2)$

Performance Test

- Take an example with 3 ordered categories ($K = 3$)
- Samples are drawn from a multinomial distribution with probabilities $\pi = (0.3, 0.5, 0.2)$
- Is asymptotic or bootstrap distribution a good approximation of the exact distribution of the statistic?
  - if we are using 95% CIs of I_α
  - coverage error rate should be close to nominal rate, 0.05

Performance Test

- Take an example with 3 ordered categories ($K = 3$)
- Samples are drawn from a multinomial distribution with probabilities $\pi = (0.3, 0.5, 0.2)$
- Is asymptotic or bootstrap distribution a good approximation of the exact distribution of the statistic?
  - if we are using 95% CIs of I_α
  - coverage error rate should be close to nominal rate, 0.05
- Check coverage error rate of CIs as sample size increases
  - $\alpha = -1, 0, 0.5, 0.99$
  - 199 bootstraps
  - 10 000 replications to compute error rates
  - $n = 20, 50, 100, 200, 500, 1000$

Estimation Methods Compared

	α	-1	0	0.5	0.99
Asymptotic B	$n = 20$	0.0606	0.0417	0.0598	0.0491
	$n = 500$	0.0523	0.0492	0.0521	0.0523
	$n = 1000$	0.0485	0.0540	0.0552	0.0549
Percentile B	$n = 20$	0.0384	0.0981	0.0912	0.1023
	$n = 500$	0.0509	0.0513	0.0552	0.0554
	$n = 1000$	0.0482	0.0556	0.0547	0.0551
Studentized B	$n = 20$	0.1275	0.0843	0.1041	0.1377
	$n = 500$	0.0518	0.0478	0.0429	0.0465
	$n = 1000$	0.0473	0.0522	0.0493	0.0503

Estimation Methods Compared

	α	-1	0	0.5	0.99
Asymptotic B	$n = 20$	0.0606	0.0417	0.0598	0.0491
	$n = 500$	0.0523	0.0492	0.0521	0.0523
	$n = 1000$	0.0485	0.0540	0.0552	0.0549
Percentile B	$n = 20$	0.0384	0.0981	0.0912	0.1023
	$n = 500$	0.0509	0.0513	0.0552	0.0554
	$n = 1000$	0.0482	0.0556	0.0547	0.0551
Studentized B	$n = 20$	0.1275	0.0843	0.1041	0.1377
	$n = 500$	0.0518	0.0478	0.0429	0.0465
	$n = 1000$	0.0473	0.0522	0.0493	0.0503

- Asymptotic CIs perform OK in finite sample

Estimation Methods Compared

	α	-1	0	0.5	0.99
Asymptotic B	$n = 20$	0.0606	0.0417	0.0598	0.0491
	$n = 500$	0.0523	0.0492	0.0521	0.0523
	$n = 1000$	0.0485	0.0540	0.0552	0.0549
Percentile B	$n = 20$	0.0384	0.0981	0.0912	0.1023
	$n = 500$	0.0509	0.0513	0.0552	0.0554
	$n = 1000$	0.0482	0.0556	0.0547	0.0551
Studentized B	$n = 20$	0.1275	0.0843	0.1041	0.1377
	$n = 500$	0.0518	0.0478	0.0429	0.0465
	$n = 1000$	0.0473	0.0522	0.0493	0.0503

- Asymptotic CIs perform OK in finite sample
- Percentile bootstrap performs well for $n > 50$

Estimation Methods Compared

	α	-1	0	0.5	0.99
Asymptotic B	$n = 20$	0.0606	0.0417	0.0598	0.0491
	$n = 500$	0.0523	0.0492	0.0521	0.0523
	$n = 1000$	0.0485	0.0540	0.0552	0.0549
Percentile B	$n = 20$	0.0384	0.0981	0.0912	0.1023
	$n = 500$	0.0509	0.0513	0.0552	0.0554
	$n = 1000$	0.0482	0.0556	0.0547	0.0551
Studentized B	$n = 20$	0.1275	0.0843	0.1041	0.1377
	$n = 500$	0.0518	0.0478	0.0429	0.0465
	$n = 1000$	0.0473	0.0522	0.0493	0.0503

- Asymptotic CIs perform OK in finite sample
- Percentile bootstrap performs well for $n > 50$
- Studentized bootstrap does not do well for small samples
- Reliable results for $\alpha = 0.99$ (index is undefined for $\alpha = 1$)

Outline

World Values Survey

World Values Survey

- Life satisfaction question:

All things considered, how satisfied are you with your life as a whole these days? Using this card on which 1 means you are "completely dissatisfied" and 10 means you are "completely satisfied" where would you put your satisfaction with your life as a whole? (code one number):

Completely dissatisfied – 1 2 3 4 5 6 7 8 9 10 – Completely satisfied

World Values Survey

- Life satisfaction question:

All things considered, how satisfied are you with your life as a whole these days? Using this card on which 1 means you are "completely dissatisfied" and 10 means you are "completely satisfied" where would you put your satisfaction with your life as a whole? (code one number):

Completely dissatisfied – 1 2 3 4 5 6 7 8 9 10 – Completely satisfied

- Health question:

All in all, how would you describe your state of health these days? Would you say it is (read out):

1 Very good, 2 Good, 3 Fair, 4 Poor.

GDP and Life satisfaction

- Cross-country comparison of life satisfaction and GDP/head
  - happiness-income paradox (Easterlin 1974, Clark and Senik 2011)
  - weak relation happiness-income internationally? (Easterlin 1995, Easterlin et al. 2010)
  - or a strong relationship? (Hagerty and Veenhoven 2003, Deaton 2008, Stevenson and Wolfers 2008a, Inglehart et al. 2008)

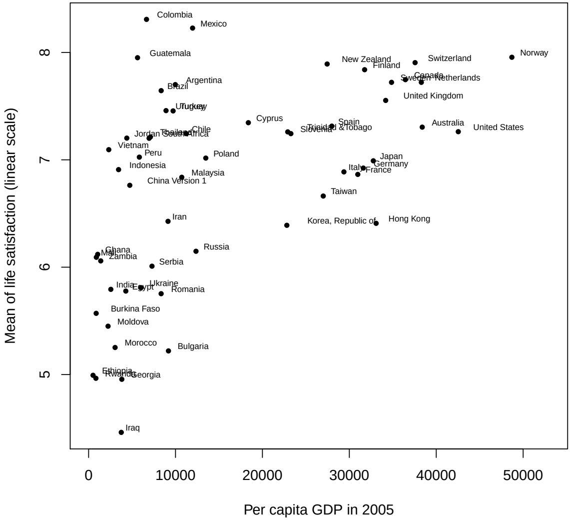
GDP and Life satisfaction

- Cross-country comparison of life satisfaction and GDP/head
  - happiness-income paradox (Easterlin 1974, Clark and Senik 2011)
  - weak relation happiness-income internationally? (Easterlin 1995, Easterlin et al. 2010)
  - or a strong relationship? (Hagerty and Veenhoven 2003, Deaton 2008, Stevenson and Wolfers 2008a, Inglehart et al. 2008)
- How should we quantify life satisfaction?
  - simple linearity of Likert scale? or exponential scale?
  - Ng (1997), Ferrer-i-Carbonell and Frijters (2004), Kristoffersen (2011)

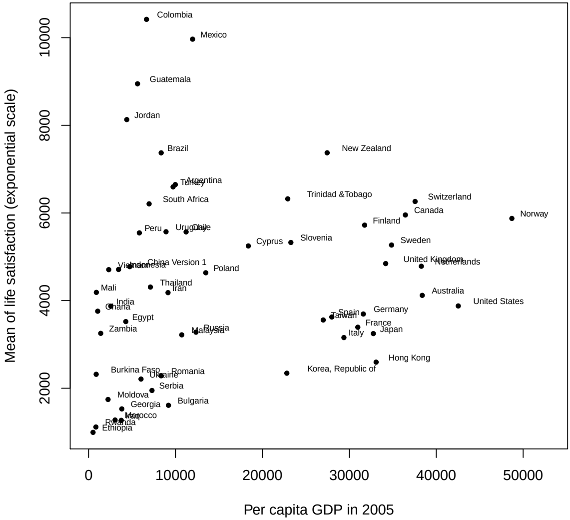
GDP and Life satisfaction

- Cross-country comparison of life satisfaction and GDP/head
  - happiness-income paradox (Easterlin 1974, Clark and Senik 2011)
  - weak relation happiness-income internationally? (Easterlin 1995, Easterlin et al. 2010)
  - or a strong relationship? (Hagerty and Veenhoven 2003, Deaton 2008, Stevenson and Wolfers 2008a, Inglehart et al. 2008)
- How should we quantify life satisfaction?
  - simple linearity of Likert scale? or exponential scale?
  - Ng (1997), Ferrer-i-Carbonell and Frijters (2004), Kristoffersen (2011)
- Is inequality of life satisfaction related to GDP/head?
  - Use I_0 and other members of the same family

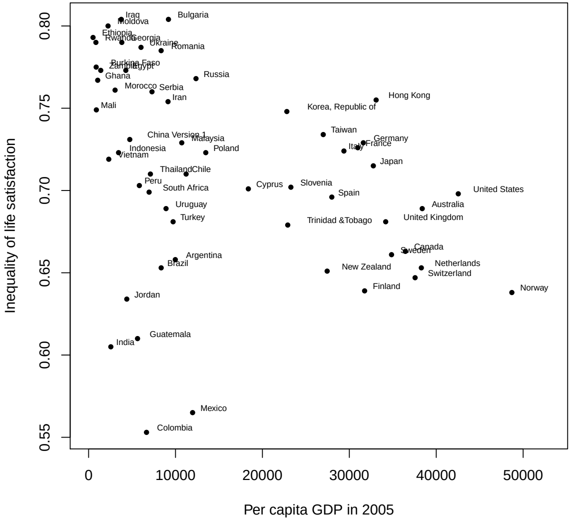
GDP and Life satisfaction (Linear)



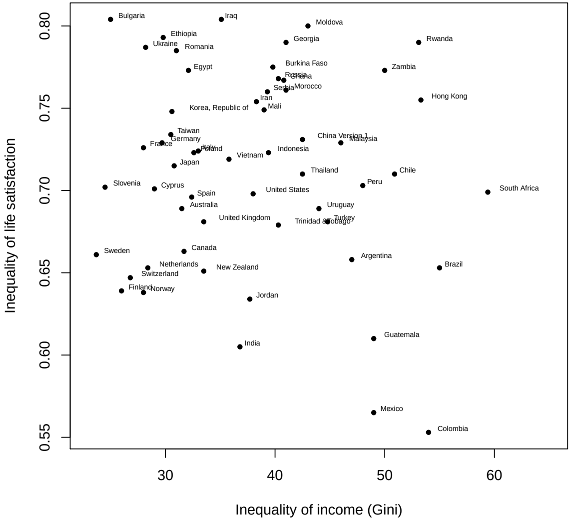
GDP and Life satisfaction (Exponential)



GDP and Inequality of Life satisfaction



Income inequality and Inequality of Life satisfaction



Health status

- Health is HRS

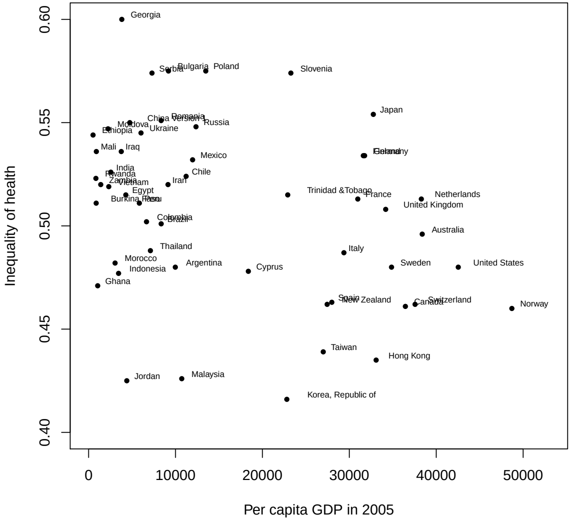
Health status

- Health is HRS
- Cross-country comparison of health and GDP
  - a significant positive relationship? (Deaton 2008)

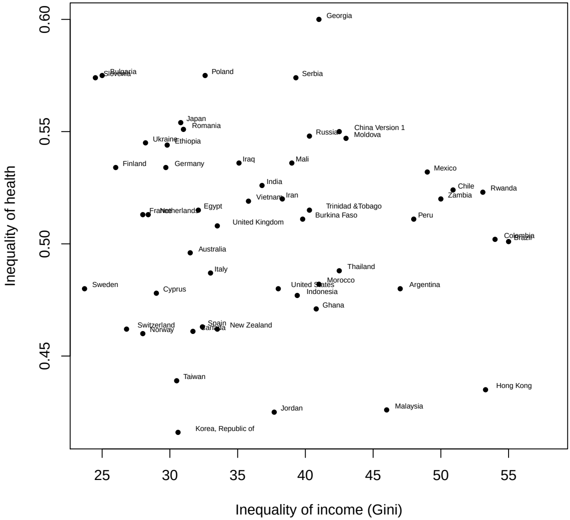
Health status

- Health is HRS
- Cross-country comparison of health and GDP
  - a significant positive relationship? (Deaton 2008)
- Cross-country comparison of inequality of health and Inequality of life satisfaction
  - use same inequality index as for life satisfaction

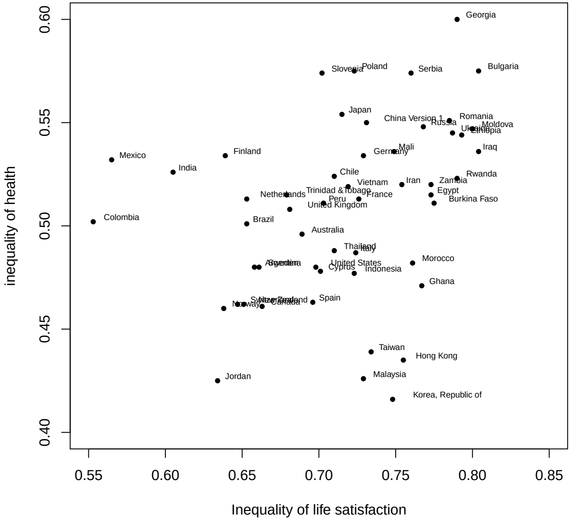
GDP and Inequality of health



Income inequality and health inequality



Inequality of life satisfaction and health inequality



Application: overview

Application: overview

- Satisfaction / GDP results sensitive to the cardinal interpretation of the answers
  - linear: positive relation below $15 000, flat after that (Layard 2003)
  - exponential: no relation

Application: overview

- Satisfaction / GDP results sensitive to the cardinal interpretation of the answers
    - linear: positive relation below $15 000, flat after that (Layard 2003)
    - exponential: no relation
- OLS estimate of I_0(life satisfaction) on the GDP per capita small and negative
    - happiness-income relationship is weak in cross-country comparisons

Application: overview

- Satisfaction / GDP results sensitive to the cardinal interpretation of the answers
  - linear: positive relation below \$15 000, flat after that (Layard 2003)
  - exponential: no relation
- OLS estimate of I_0(life satisfaction) on the GDP per capita small and negative
  - happiness-income relationship is weak in cross-country comparisons
- No clear relationship between I_0(health) on GDP per capita

Application: overview

- Satisfaction / GDP results sensitive to the cardinal interpretation of the answers
  - linear: positive relation below \$15 000, flat after that (Layard 2003)
  - exponential: no relation
- OLS estimate of I_0(life satisfaction) on the GDP per capita small and negative
  - happiness-income relationship is weak in cross-country comparisons
- No clear relationship between I_0(health) on GDP per capita
- OLS estimate of I_0(health) on I_0(life satisfaction) produces a slope coefficient not significantly different from zero
  - health-life satisfaction relationship is not significant

Summary

- Inequality with ordinal data is a widespread phenomenon

Summary

- Inequality with ordinal data is a widespread phenomenon
- Conventional I-measures may make no sense

Summary

- Inequality with ordinal data is a widespread phenomenon
- Conventional I-measures may make no sense
- Cowell and Flachaire (2014) approach:
  - separates out the issue of status from that of inequality-aggregation
  - allows you to choose "reference status"
  - gives a family of measures

Summary

- Inequality with ordinal data is a widespread phenomenon
- Conventional *I*-measures may make no sense
- Cowell and Flachaire (2014) approach:
  - separates out the issue of status from that of inequality-aggregation
  - allows you to choose “reference status”
  - gives a family of measures
- Nice properties empirically

Summary_

Outline

Median: definition and in our cases

- $\text{med}(\mathbf{s})$ defined as $e \in S$ such that $\#(s_i \leq e) \geq \frac{n}{2}, \#(s_i \geq e) \geq \frac{n}{2}$

Median: definition and in our cases

- med(**s**) defined as $e \in S$ such that $\#(s_i \leq e) \geq \frac{n}{2}, \#(s_i \geq e) \geq \frac{n}{2}$

	Case 0		Case 1		Case 2		Case 3	
	n_k	s_i	n_k	s_i	n_k	s_i	n_k	s_i
B	0		25	1	0		25	1
E	50	1	25	$3/4$	50	1	25	$3/4$
G	25	$1/2$	25	$1/2$	50	$1/2$	50	$1/2$
N	25	$1/4$	25	$1/4$	0		0	
$M(\mathbf{s})$	$[1/2, 1)$		$[1/2, 3/4)$		$[1/2, 1)$		$[1/2, 3/4)$	

- med(**s**) could be any value in interval $M(\mathbf{s})$

Median example 1

- Three ordered categories
- Same proportion of individuals in each category

Median example 1

- Three ordered categories
- Same proportion of individuals in each category
- The status vector is $\mathbf{s} = ({}^{1}/_{3}, {}^{2}/_{3}, 1)$

Median example 1

- Three ordered categories
- Same proportion of individuals in each category
- The status vector is $\mathbf{s} = (1/3, 2/3, 1)$
- conventional definition is $\mathrm{med}(\mathbf{s}) = m := 2/3$:
  - $2/3$ of the population has a status less or equal to m
  - $2/3$ of the population has a status greater than or equal to m

Median example 1

- Three ordered categories
- Same proportion of individuals in each category
- The status vector is $\mathbf{s} = ({}^{1}\!/_{3}, {}^{2}\!/_{3}, 1)$
- conventional definition is $\text{med}(\mathbf{s}) = m := {}^{2}\!/_{3}$:
  - ${}^{2}\!/_{3}$ of the population has a status less or equal to m
  - ${}^{2}\!/_{3}$ of the population has a status greater than or equal to m
- Median as “half-way” point is misleading

Median example 2

- Two ordered categories (B better than A)
- Three distributions

1. $n_{\mathrm{A}} = 500,\ n_{\mathrm{B}} = 500$
2. $n_{\mathrm{A}} = 499,\ n_{\mathrm{B}} = 501$
3. $n_{\mathrm{A}} = 999,\ n_{\mathrm{B}} = 1$

Median example 2

- Two ordered categories (B better than A)
- Three distributions
  1. $n_A = 500,\ n_B = 500$
  2. $n_A = 499,\ n_B = 501$
  3. $n_A = 999,\ n_B = 1$
- Status and median in each case:
  1. $\mathbf{s} = (0.5, 1)$, $\text{med}(\mathbf{s}) = 0.5$
  2. $\mathbf{s} = (0.499, 1)$, $\text{med}(\mathbf{s}) = 1$
  3. $\mathbf{s} = (0.999, 1)$, $\text{med}(\mathbf{s}) = 0.999$

Median example 2

- Two ordered categories (B better than A)
- Three distributions

1. $n_{\mathrm{A}} = 500,\ n_{\mathrm{B}} = 500$
2. $n_{\mathrm{A}} = 499,\ n_{\mathrm{B}} = 501$
3. $n_{\mathrm{A}} = 999,\ n_{\mathrm{B}} = 1$

- Status and median in each case:

1. $\mathbf{s} = (0.5, 1)$, $\mathrm{med}(\mathbf{s}) = 0.5$
2. $\mathbf{s} = (0.499, 1)$, $\mathrm{med}(\mathbf{s}) = 1$
3. $\mathbf{s} = (0.999, 1)$, $\mathrm{med}(\mathbf{s}) = 0.999$

- Compare:
  - distributions 1 and 2 have very different medians
  - distributions 2 and 3 have almost the same median!

Median example 2

- Two ordered categories (B better than A)
- Three distributions

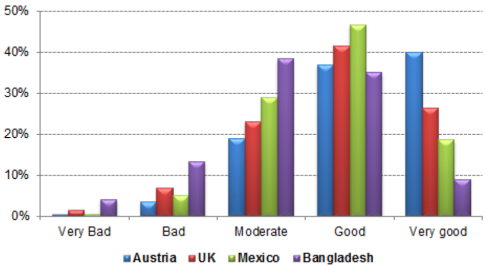
1. $n_{\mathrm{A}} = 500,\ n_{\mathrm{B}} = 500$
2. $n_{\mathrm{A}} = 499,\ n_{\mathrm{B}} = 501$
3. $n_{\mathrm{A}} = 999,\ n_{\mathrm{B}} = 1$

- Status and median in each case:

1. $\mathbf{s} = (0.5, 1)$, $\mathrm{med}(\mathbf{s}) = 0.5$
2. $\mathbf{s} = (0.499, 1)$, $\mathrm{med}(\mathbf{s}) = 1$
3. $\mathbf{s} = (0.999, 1)$, $\mathrm{med}(\mathbf{s}) = 0.999$

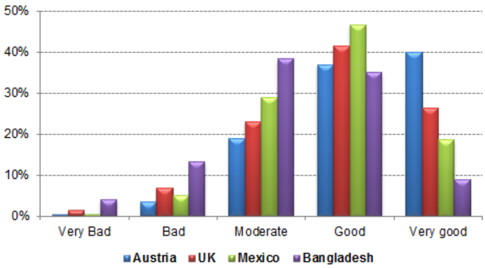
- Compare:
  - distributions 1 and 2 have very different medians
  - distributions 2 and 3 have almost the same median!

SRH Inequality: Gini (median norm'd)



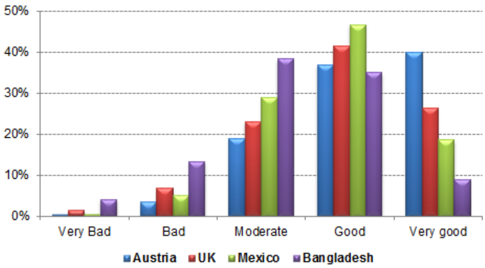
	At	UK	Mx	BD	
(1,2,3,4,5)	0.107	0.135	0.123	0.140	(BD,UK,Mx,At)*

SRH Inequality: Gini (median norm'd)



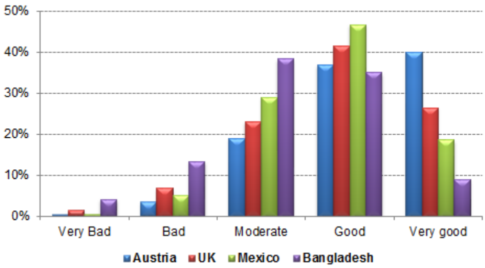
	At	UK	Mx	BD	
(1,2,3,4,5)	0.107	0.135	0.123	0.140	(BD,UK,Mx,At)*
(1,2,3,4,1000)	0.006	0.011	0.017	0.029	(BD,Mx,UK,At)*

SRH Inequality: Gini (median norm’d)



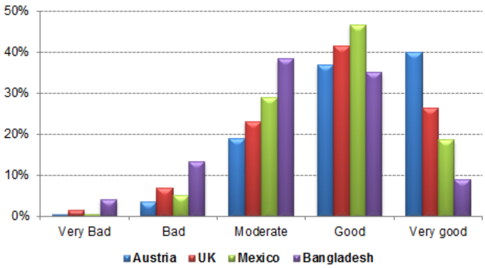
	At	UK	Mx	BD	
(1,2,3,4,5)	0.107	0.135	0.123	0.140	(BD,UK,Mx,At)*
(1,2,3,4,1000)	0.006	0.011	0.017	0.029	(BD,Mx,UK,At)*
(-1000,2,3,4,5)	7.39	-0.315	-1.844	-0.188	(At,Mx,UK,BD)

SRH Inequality: C of V (median norm'd)



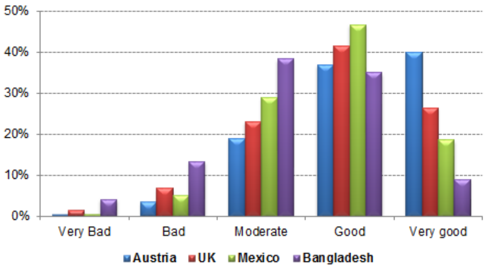
		At	UK	Mx	BD	
(1,2,3,4,5)		0.202	0.253	0.232	0.260	(BD,UK,Mx,At) *

SRH Inequality: C of V (median norm'd)



	At	UK	Mx	BD	
(1,2,3,4,5)	0.202	0.253	0.232	0.260	(BD,UK,Mx,At) *
(1,2,3,4,1000)	0.012	0.024	0.044	0.101	(BD,Mx,UK,At)*

SRH Inequality: C of V (median norm'd)



	At	UK	Mx	BD	
(1,2,3,4,5)	0.202	0.253	0.232	0.260	(BD,UK,Mx,At) *
(1,2,3,4,1000)	0.012	0.024	0.044	0.101	(BD,Mx,UK,At)*
(-1000,2,3,4,5)	2276	-4.39	-87.2	-0.42	(At,BD,UK,Mx)

Outline

Proof of Theorem 1

- Two cases to consider
  - data are categorical: S is set of non-negative rational numbers, $\mathbb{Q}_+$.
  - data have cardinal significance: S can be taken as an interval in $\mathbb{R}$.
- In either case $(S, +, \succ)$ forms a strictly ordered group (Krantz 1964, Luce and Tukey 1964, Wakker 1988)
- From Theorem 5.3 of Fishburn (1970) Axioms jointly imply that, for a given e, $\succeq$ is representable by a continuous function $S^{n+1} \to \mathbb{R}$:$\sum_{i=1}^{n} d_i(s_i, e), \forall (\mathbf{s}, e) \in S^{n+1}$where, for each i, $d_i : S \to \mathbb{R}$ is a continuous function.
- By monotonicity this is increasing in s_i if $s_i > e$ and *vice versa*.
- By anonymity the functions d_i must all be identical
- ordering $\succeq$ is also representable any monotonic transform

"Maximum inequality"

- Take the case where status is downward-looking and peer-inclusive
- Suppose that the status of each member of category k is s
- If a person is promoted from category k to category $k+1$
  - status increases to $s + n_{k+1}/n$
  - status of each of the remaining $n_k - 1$ members of category k falls to $s - 1/n$.
- The resulting change in inequality iproportional to $\left[d\left(s+\frac{n_{k+1}}{n},e\right)-d\left(s,e\right)\right]+\left[n_k-1\right]\left[d\left(s-\frac{1}{n},e\right)-d\left(s,e\right)\right]$
- If d is differentiable then this expression is approximately $d'\left(s,e\right)\frac{n_{k+1}}{n}-\frac{n_k-1}{n}d'\left(s,e\right)$
  - which equals $\frac{1}{n}d'\left(s,e\right)\left[n_{k+1}-n_k+1\right]$.
- If $s < e$ then monotonicity implies $d'\left(s,e\right) < 0$
  - the change in inequality is negative if $n_{k+1} \geq n_k$.

Outline

Dispersion

- Model: $\text{LifeSatisf}_i = \alpha + \beta\,\text{GDP}_i + \varepsilon_i, \qquad \varepsilon_i \sim N(0, \sigma_i^2)$
  - β is a significant coefficient and R^2 is large
  - strong (linear) relationship between LifeSatisf and GDP
- If LifeSatisf equation is homoskedastic:
  - no relationship between GDP and the dispersion of LifeSatisf
  - whatever is GDP, the dispersion of LifeSatisf is the same
- If LifeSatisf equation heteroskedastic dispersion of LifeSatisf may or may not be related to GDP
  - the form of the heteroskedasticity cannot be deduced from the relationship between the dependent variable and the covariate.
- If every i has different GDP, σ_i^2 measures the dispersion of LifeSatisf for i
- taking the measure I_0 as a measure of dispersion, the same reasoning applies

Bibliography I

Abul Naga, R. H. and T. Yalcin (2008). Inequality measurement for ordered response health data. *Journal of Health Economics 27*, 1614–1625.

Abul Naga, R. H. and T. Yalcin (2010). Median independent inequality orderings. Technical report, University of Aberdeen Business School.

Allison, R. A. and J. E. Foster (2004). Measuring health inequality using qualitative data. *Journal of Health Economics 23*, 505–552.

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory 2*, 244–263.

Clark, A. E. and C. Senik (2011). Will GDP growth increase happiness in developing countries? In J. Slemrod (Ed.), *Measure For Measure: How well do we Measure Development?* AFD Publications.

Cowell, F. A. and E. Flachaire (2014). Inequality with ordinal data. Public Economics Programme Discussion Paper 16, London School of Economics, http://darp.lse.ac.uk/pdf/IneqOrdinal.pdf.

Dalton, H. (1920). Measurement of the inequality of incomes. *The Economic Journal 30*, 348–361.

de Barros, R. P., F. Ferreira, J. Chanduvi, and J. Vega (2008). *Measuring Inequality of Opportunities in Latin America and the Caribbean*. Palgrave Macmillan.

Deaton, A. (2008). Income, health and well-being around the world: Evidence from the Gallup World Poll. *Journal of Economic Perspectives 22*, 53–72.

Easterlin, R. A. (1974). Does economic growth improve the human lot? Some empirical evidence. In P. A. David and M. W. Reder (Eds.), *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz*. New York: Academic Press.

Bibliography II

Easterlin, R. A. (1995). Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior & Organization 27*, 35–47.

Easterlin, R. A., L. Angelescu McVey, M. Switek, O. Sawangfa, and J. Smith Zweig (2010). The happiness-income paradox revisited. *Proceedings of the National Academy of Sciences of the United States of America 107*, 22463–22468.

Ferrer-i-Carbonell, A. and P. Frijters (2004). How important is methodology for the estimates of the determinants of happiness? *The Economic Journal 114*, 641–659.

Fishburn, P. C. (1970). *Utility Theory for Decision Making*. New York: John Wiley.

Hagerty, M. R. and R. Veenhoven (2003). Wealth and happiness revisited: Growing wealth of nations does go with greater happiness. *Social Indicators Research 64*, 1–27.

Inglehart, R., R. Foa, C. Peterson, and C. Welzel (2008). Development, freedom, and rising happiness: A global perspective (1981-2007). *Perspectives on Psychological Science 3*, 264–285.

Krantz, D. H. (1964). Conjoint measurement: The luce-tukey axiomatization and some extensions. *Journal of Mathematical Psychology 1*(2), 248 – 277.

Kristoffersen, I. (2011). The subjective wellbeing scale: How reasonable is the cardinality assumption? Discussion Paper 15, University of Western Australia Department of Economics.

Layard, R. (2003). Happiness: Has social science a clue. Lionel Robbins Memorial Lectures 2002/3, London School of Economics, march 3-5. http://cep.lse.ac.uk/events/lectures/layard/RL030303.pdf.

Luce, R. and J. W. Tukey (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology 1*(1), 1 – 27.

Bibliography III

Ng, Y. K. (1997). A case for happiness, cardinalism, and interpersonal comparability. *The Economic Journal 107*, 1848–58.

Oswald, A. J. and S. Wu (2011, November). Well-being across America. *The Review of Economics and Statistics 93*(4), 1118–1134.

Stevenson, B. and J. Wolfers (2008a). Economic growth and subjective well-being: Reassessing the Easterlin paradox. NBER working paper no. 14282.

Stevenson, B. and J. Wolfers (2008b). Happiness inequality in the United States. *The Journal of Legal Studies 37*, S33–S79.

Van Doorslaer, E. and A. M. Jones (2003). Inequalities in self-reported health: Validation of a new approach to measurement. *Journal of Health Economics 22*, 61–87.

Wakker, P. (1988). The algebraic versus the topological approach to additive representations. *Journal of Mathematical Psychology 32*(4), 421 – 435.

Yang, Y. (2008). Social inequalities in happiness in the United States, 1972 to 2004: An age-period-cohort analysis. *American Sociological Review 73*, 204–226.

Yitzhaki, S. (1979). Relative deprivation and the Gini coefficient. *Quarterly Journal of Economics 93*, 321–324.

Zheng, B. (2011). A new approach to measure socioeconomic inequality in health. *Journal Of Economic Inequality 9*, 555–577.