Summer School on Capability and Multidimensional Poverty

11-20 September 2010

Amman, Jordan

OPHI

Oxford Poverty & Human Development Initiative

University of Oxford

www.ophi.org.uk

Inequality-Adjusted HDI

Sabina Alkire (OPHI)

James Foster (GW, OPHI)

Working Paper 37

Human Development Index

- Motivation
  - Go beyond per capita income as a wellbeing measure
    - Ends as well as means
  - Broaden space
    - First form of heterogeneity (F. Bourguignon)
  - Practical indicator
    - Substantial coverage with existing data
    - Easy to understand

Human Development Index

Issues

(i) Choice of dimensions (variables)

income (GDP per capita – we use GNI per capita)

education (literacy/enrolment we use Yrs Ed/school LE)

health (life expectancy - LE)

Why only these?

(ii) Measurability of variables (cardinal or ordinal)

assumes cardinal (at least interval scale)

How can you justify?

(iii) Comparability of variables(full, partial or not at all)

full after normalizing to a common range [0,1]

How does empirical become normatively relevant?

(iv) Aggregation and weighting (general functions?)

mean of means

What about second form of heterogeneity? Inequality

Alkire & Foster 2010 - IHDI

Theory

- Address issue (iv) inequality using FLS (2005)
  - **Assume issues (i-iii) solved**
  - "after appropriate transformations, all variables are measured using a ratio scale in such a way that levels are comparable across dimensions"
  - Provide plausible calibration method
- Construct IHDI
  - Based on Atkinson's ede instead of arithmetic mean
  - Focus on H_1 using geometric mean
  - And H_1* suppressing dimensional inequality
- Interpretation
  - IHDI and potential IHDI
  - Inequality adjustment used below in estimation
- Properties
  - H_1 satisfies usual properties
  - Invariance properties via geometric mean used below in ensuring robustness to calibration choices

Implementation

- Revisit measurement assumptions (ii-iii)
  - Calibrating variables
- Estimating indices
  - Potential IHDI H_1^*
    - Uses aggregate data arithmetic means
    - Combine using geometric mean
  - IHDI H_1
    - Geometric means unavailable for aggregate data
    - Use estimates of Atkinson's inequality measure to adjust mean
    - Combine using geometric mean
- Example

General means – different story

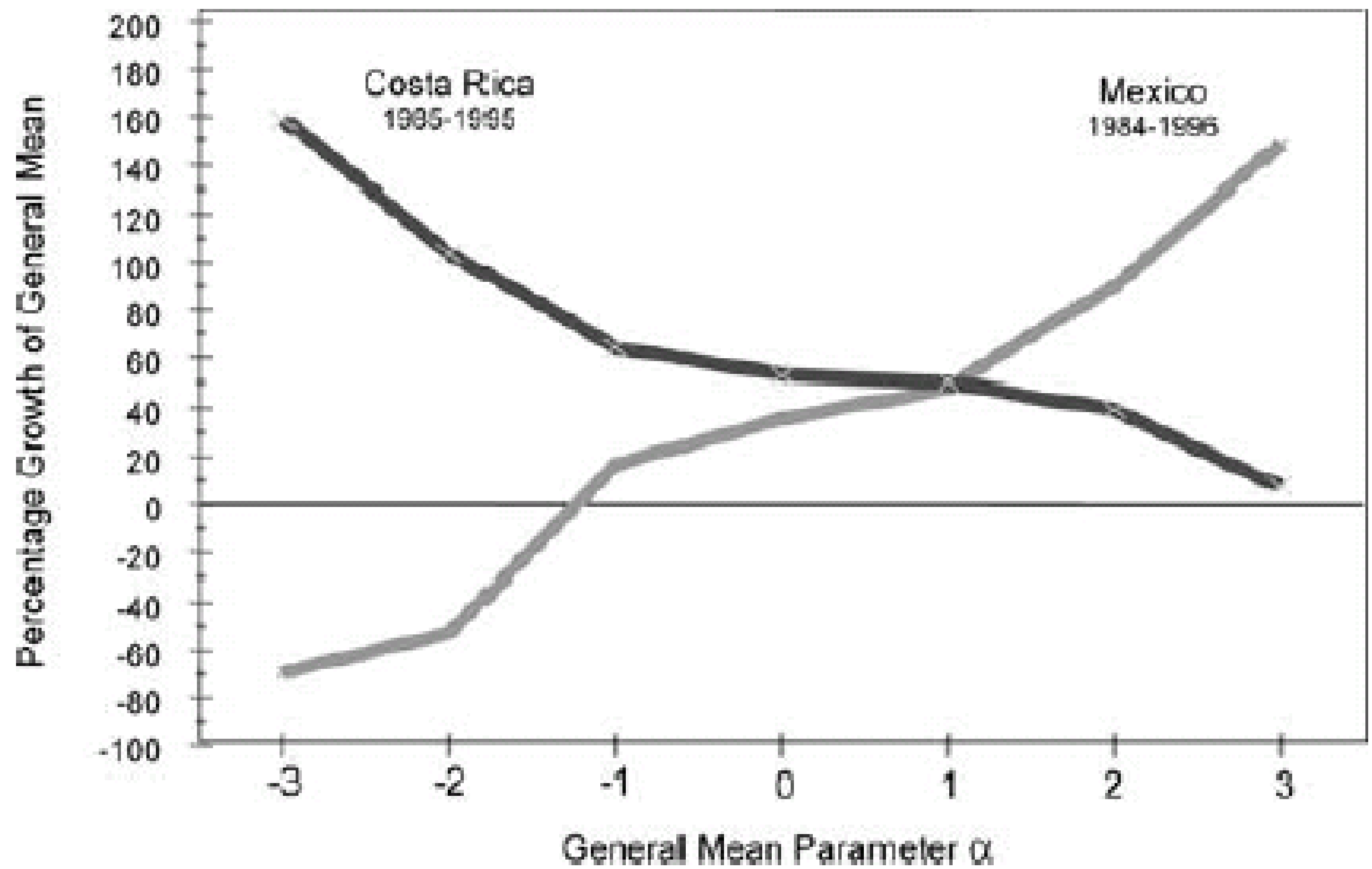


FIGURE 2

GROWTH IN THE INCOME STANDARDS OF COSTA RICA AND MEXICO

Review of Foster Lopez-Calva Szekely (FLS) 2005 in *J of Human Development*

Notation

- x distribution of income
- y distribution of education
- z distribution of health
- D matrix of achievements

HDI

$H(D) = \mu[\mu(x), \mu(y), \mu(z)]$

Measure of average achievement

Equally distributed equivalent

- Assuming welfare has form
- $W(D) = \Sigma_i \Sigma j\ u(d_{ij})$ with u linear

HDI

$$H(D) = \mu[\mu(x),\mu(y),\mu(z)]$$

Properties

- symmetry in dimensions
- symmetry in people
- replication invariance
- normalization
- linear homogeneity
- monotonicity
- subgroup consistency

Problem

- Like per capita GNI, ignores inequality

Gini-adjusted HDI

Anand & Sen (1993) and Hicks (1997)

Use Sen welfare index to include inequality within dimensions

$S(x) = \mu(x)[1-G(x)]$	income
$S(y) = \mu(y)[1-G(y)]$	education
$S(z) = \mu(z)[1-G(z)]$	health

Note Mean achievement is discounted by inequality

Gini-adjusted HDI

$H_G(D) = \mu[S(x),S(y),S(z)]$

Properties

Symmetry in dimensions, symmetry in people, replication invariance, normalization, linear homogeneity, monotonicity

Violates subgroup consistency

$H_G(D)$ rises $H_G(D')$ rises $H_G(D;D')$ falls

Gain inequality sensitivity - but at some cost

Not applicable to regional analysis

Also not "path independent"

Results depend on order of aggregation

- people then dimensions vs. dimensions then people

Note Culprit is Gini in Sen welfare measure

Alternatives? Foster, Lopez Calva, Szekely (2005)

Inequality-adjusted HDI (IHDI)

Recall

Equally distributed equivalent income (ede)

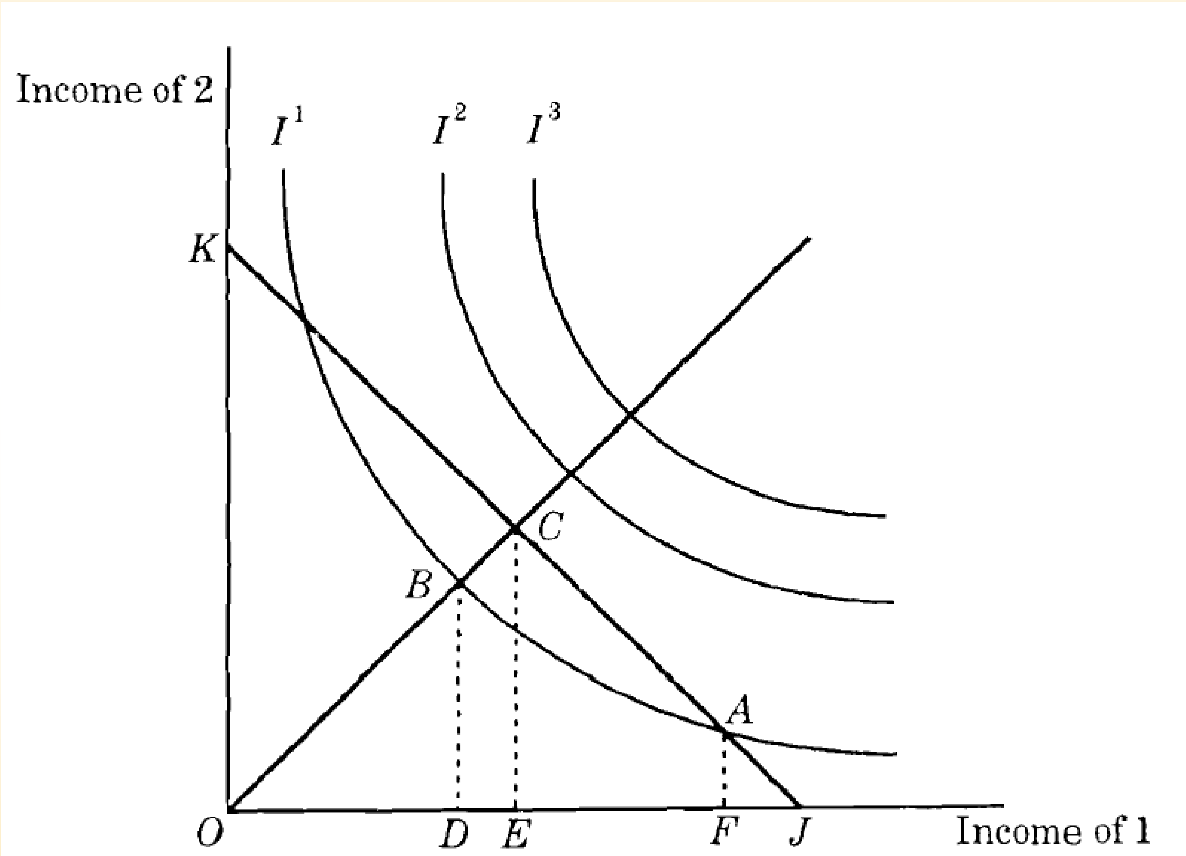
The income level which, if assigned to all individuals, produces the same social welfare as the observed distribution.

Note

For any W(x), the associated ede e(x) always ranks distribution the same way as W(x)

The ede e(x) is a welfare function

Equally Distributed Equivalent (ede)



A = initial income distribution

Three social welfare levels, I^1, I^2, I^3

Find the following:

Total income

Mean Income

Set of all possible distributions

Equally distributed equivalent income

Atkinson's ede

$$e_\alpha(x) = \begin{cases} \left[\frac{1}{n}\sum_{i=1}^{n} x_i^\alpha\right]^{1/\alpha} & \alpha \le 1, \alpha \ne 0 \\ \prod_{i=1}^{n} x_i^{1/n} & \alpha = 0 \end{cases}$$

$e_\alpha(x) = \mu_\alpha(x)$ general mean

$A_\varepsilon(x) = [\mu(x) - \mu_\alpha(x)]/\mu(x)$ inequality

$\mu_\alpha(x) = \mu(x)[1 - A_\varepsilon(x)]$ discounted for ineq

Inequality-adjusted HDI (IHDI)

$H_\varepsilon(D) = \mu_\alpha(D)$ for $\varepsilon > 0$, $\alpha = 1 - \varepsilon$

- general mean applied to matrix
- ede achievement level

$\varepsilon = 0$ $H_0 = \mu[D]$ usual HDI

$\varepsilon = 1$ $H_1 = \mu_0 [D]$

based on geometric mean $g = \mu_0$

sensitive to inequality

$\varepsilon = 2$ $H_2 = \mu_{-1} [D]$

based on harmonic mean μ_{-1}

even more sensitive

Inequality-adjusted HDI (IHDI)

$$H_\varepsilon(D) = H_0(D)[1 - A_\varepsilon(D)]$$

Note

- Inequality adjusted
- Which inequalities?
  - Both within dimensions
  - And across dimensions

Interpretation

- H_0 is highest possible level of H_ε when one can freely transfer achievements across achievements and dimension
- H_ε indicates the actual IHDI

IHDI

Properties

- Symmetry in dimensions, symmetry in people, replication invariance, normalization, linear homogeneity, monotonicity
- Subgroup consistency

IHDI

Alternative definitions

$H_\varepsilon(D) = \mu_\alpha[\mu_\alpha(x), \mu_\alpha(y), \mu_\alpha(z)]$ aggregate within dimensions then across dimensions

$H_\varepsilon(D) = \mu_\alpha[\mu_\alpha(d_1),\ldots,\mu_\alpha(d_m)]$ aggregate at individual level then across persons

Path independence

- Conceptual
- Empirical – need only $A_\varepsilon(x)$, $A_\varepsilon(y)$, $A_\varepsilon(z)$ estimates

Note: Dimensions are not perfect substitutes

- Measure of complementarity: ε
- Sensitive to uneven growth
- Not sensitive to correlations!
  - See work by Seth (2010)

Family of Human Development Indices for Mexican States

Ranking	State	H index e=0	State	H index e=3	Ranking
1	Chiapas	0.5735	Oaxaca	0.3654	1
2	Oaxaca	0.5881	Chiapas	0.3797	2
3	Guerrero	0.5968	Guerrero	0.3995	3
4	Veracruz	0.6168	Veracruz	0.4337	4
5	Puebla	0.6232	Zacatecas	0.4401	5
6	Yucat‡n	0.6239	Yucat‡n	0.4497	6
7	Michoac‡n	0.6363	Michoac‡n	0.4509	7
8	San Luis Potos’	0.6370	Puebla	0.4545	8
9	Hidalgo	0.6449	San Luis Potos’	0.4641	9
10	Zacatecas	0.6482	Durango	0.4708	10
11	Guanajuato	0.6546	Tlaxcala	0.4747	11
12	Tlaxcala	0.6600	Hidalgo	0.4784	12
13	Durango	0.6608	Nayarit	0.4898	13
14	Quer˘taro	0.6637	Guanajuato	0.4937	14
15	Nayarit	0.6638	Chihuahua	0.5069	15
16	Tabasco	0.6646	Tabasco	0.5094	16
17	Morelos	0.6691	Morelos	0.5139	17
18	Campeche	0.6734	Quer˘taro	0.5146	18
19	Chihuahua	0.6739	M˘xico	0.5185	19
20	Tamaulipas	0.6752	Jalisco	0.5246	20
21	Jalisco	0.6772	Sonora	0.5256	21
22	Quintana Roo	0.6798	Tamaulipas	0.5280	22
23	Sinaloa	0.6817	Colima	0.5428	23
24	M˘xico	0.6824	Quintana Roo	0.5438	24
25	Sonora	0.6853	Sinaloa	0.5472	25
26	Colima	0.6884	Campeche	0.5473	26
27	Coahuila	0.6957	Coahuila	0.5637	27
28	Aguascalientes	0.7001	Nuevo Le—n	0.5783	28
29	Nuevo Le—n	0.7021	Baja California Sur	0.5787	29
30	Baja California Sur	0.7038	Aguascalientes	0.5811	30
31	Baja California	0.7176	Baja California	0.6150	31
32	Distrito Federal	0.7403	Distrito Federal	0.6376	32

Source: Authors« calculations using the Mexican Census 2000 sample.

IHDI

Will focus on H_1 as our key IHDI

$$H_1(D) = g(D) = g[g(x), g(y), g(z)]$$
$$= g[g(x), g(y), g(z)]$$
$$= g[\ \mu(x)[1-A(x)], \mu(y)[1-A(y)], \mu(z)[1-A(z)]\ ]$$

where $A(x) = 1 - g(x)/\mu(x)$ etc is a transform of Theil's second

IHDI

- Interesting measurement properties
  - Individual Scale Invariance
    - Changing scale of a single variable preserves ranks
    - And percentage changes
  - Independence of Standardized Values
    - Normalize to one country's achievements
    - Preserves ranks and percentage changes
  - Consistency over Time

IHDI

A second index H_1^*

$$H_1^*(D) = H_1(D^*) = g[\mu(x), \mu(y), \mu(z)]$$

Contrast with

$$H_1(D) = g[\mu(x)[1-A(x)], \mu(y)[1-A(y)], \mu(z)[1-A(z)]]$$

Idea

H_1^* is highest possible level of H_ε when one can freely transfer achievements across achievements

= Potential H_1

IHDI

A reinterpretation

$\ln H_1(D) = \mu [\ln g(x), \ln g(y), \ln g(z)]$

$\ln H_1^*(D) = \mu [\ln \mu(x), \ln \mu(y), \ln \mu(z)]$

Traditional HDI

$H_T(D) = \mu [\ln \mu(x), \mu(y), \mu(z)]$

Constructing Variables: Aggregate

Lower cutoffs

- natural zero point in the raw data
- point at which life unsustainable

- Education:
  - No schooling or prospects of schooling
- Command over resources:
  - No income
- Health
  - Zero years life expectancy
  - Lowest group LE that sustains species (15-20)

Constructing Variables: Aggregate

Upper cutoffs

- establish values of comparable fulfilment
- upper cutoff can be set as increment of these
- common method: highest observed value
- normative method: sufficiency cut-off

- Education – highest observed
- Life Expectancy – highest observed (83.2)
- Income – 1997 $40,000 = $51,200

Constructing Variables: Aggregate

Normalized variable:

x, z [Value/Max]

y [(Value1/Max1)+(Value2/Max2)]/2

Estimating Inequality from hh data

Ideally use same data as HDI

Impossible

Response: estimate (1-A) from related variables

GNI	Income/Consumption
Education	Years of Schooling 25+
LE	LE by cohort

Estimating Inequality from hh data

Command over Resources

Adjust distribution

Use household income *and* consumption per capita

Trim upper 0.5% of sample

Identify $X_{0.5}$ at 0.5% of all positive values.

Replace all entries having values $<X_{0.5}$ with $X_{0.5}$

Estimate $\hat{g}_x$ ***from the adjusted distribution***

(concerns: mix of i/c, data quality, 0 replacements)

Estimating Inequality from hh data

Education

Adjust distribution

Use years of schooling for adults 25+ years

Convert levels of schooling into years if required

Add one year of education to all values (replace 0s)

Estimate $\hat{g}_x$ ***from the adjusted distribution***

(concerns: data quality & accuracy, 0 replacements

Estimating Inequality from hh data

Health

Use Life Tables of UN Population Division

Construct cohorts (0,1), [1…5), [5…10)…[85, +).

Construct interval-average age at death (LE)

Weight cohorts

Estimate $\hat{g}_z$ ***across cohorts***

(concerns: modelling, group data, no disability or morbidity)

Data:

Aggregate HDI: plus mean years of schooling and school life expectancy

Atkinson Inequality Adjustments:

Income: World Bank compiled dataset, LIS, EU-SILC, WIDER, DHS (?)

Education: World Bank compiled dataset, LIS, EU-SILC (potential: Barro-Lee 2010), DHS

Health: Life Tables, UN Population Division 08

Consider the achievements of two countries:

		Indicators						
		Life expectancy at birth		**Mean years of schooling**		**Expected years of schooling**		**GNI per capita**
		(years)		(years)		(years)		(PPP US$)
Norway		81.0		12.6		17.3		58,810
Haiti		61.7		4.9		6.8	g	949

Define Bounds

	Upper	Lower
Income	51,200 $	0
Life Expectancy	83.2 yrs	0*
Mean Yrs School	12.6 yrs	0
School Life Expect	20.5 yrs	0

Create Normalized Indices

$$\mu_x = \frac{I}{51200}$$

$$\mu_z = \frac{H}{83.2}$$

$$\mu_y = \left(\frac{\left(\frac{E_m}{12.6}\right) + \left(\frac{E_{sl}}{20.5}\right)}{2} \right)$$

		Indicators						
		Life expectancy at birth		Mean years of schooling		Expected years of schooling		GNI per capita
		(years)		(years)		(years)		(PPP US$)
Norw ay		81.0		12.6		17.3		58,810
Haiti		61.7		4.9		6.8	g	949

Haiti

$$\mu_x = \frac{949}{51200} = 0.019$$

$$\mu_z = \frac{60.7}{83.2} = 0.741$$

$$\mu_y = \left(\frac{\left(\frac{4.9}{12.6}\right) + \left(\frac{6.8}{20.5}\right)}{2} \right) = 0.361$$

Normalized Indices

	Dimension Indices - HDI				
	Health		Education		Income
Norway	0.974		0.924		1.000
Haiti	0.741		0.361		0.019

Create H_1

1. Adjust normalized means by estimated inequality

$$\hat{g}_x = \mu_x(1 - \hat{A}_x)$$

$$\hat{g}_y = \mu_y(1 - \hat{A}_y)$$

$$\hat{g}_z = \mu_z(1 - \hat{A}_z)$$

1. Adjust normalized means by estimated inequality

$Haiti:$

$$\hat{g}_x = 0.019 \times 0.748 = 0.014$$

$$\hat{g}_z = 0.741 \times .672 = 0.498$$

$$\hat{g}_y = 0.361 \times .587 = 0.212$$

$Norway:$

$$\hat{g}_x = 1.0 \times .865 = 0.865$$

$$\hat{g}_z = 0.974 \times .961 = 0.935$$

$$\hat{g}_y = 0.924 \times .980 = 0.905$$

Ratio of geometric to arithmetic means -				
		Health	Income	Education
		g/a	g/a	g/a
Norw ay		0.9606	0.8650	0.9800
Haiti		0.6716	0.7481	0.5868

Create H_1

2. Take geometric mean of adjusted normalized variables

$$H_1 = \left(\hat{g}_x \times \hat{g}_y \times \hat{g}_z\right)^{1/3}$$

Create H_1

2. Take geometric mean of adjusted normalized variables

$$Norway:$$

$$H_1 = \left(0.865 \times 0.935 \times 0.965\right)^{1/3} = 0.901$$

$$Haiti$$

$$H_1 = \left(0.014 \times 0.498 \times 0.212\right)^{1/3} = 0.114$$

- H_1 is the Inequality Adjusted HDI
- What if we removed the inequality within each dimension?
- We would have H_1*
- H_1* reflects the *potential* HDI

Create H_1

$$H_1^* = \left(\mu_x \times \mu_y \times \mu_z\right)^{1/3}$$

Create H_1

	Dimension Indices - HDI				
	Health		Education		Income
Norway	0.974		0.924		1.000
Haiti	0.741		0.361		0.019

$$H_1^* = \left(\mu_x \times \mu_y \times \mu_z\right)^{1/3}$$

$$H_{Norway}^* = \left(1.0 \times 0.974 \times 0.924\right)^{1/3} = 0.965$$

$$H_{Haiti}^* = \left(0.019 \times 0.741 \times 0.361\right)^{1/3} = 0.171$$

Compare H_1 and H_1^*

Norway: $H_1 = 0.901$ $H_1^* = 0.965$

Haiti: $H_1 = 0.114$ $H_1^* = 0.171$

The percentage change represents the inequality adjustment: Haiti's actual IHDI is 66% of potential; Norway is 93%.

Some Explorations:

1. Use a higher 'natural zero' lower bound for health (15)
2. Use a higher GNI upper bound (observed value)
3. Enter log GNI (as in previous HDI)
4. Aggregate the two aggregate educational variables by taking their geometric mean, and normalizing by the maximum resulting value. $y' = (E_M E_{SLE})^{0.5}/(MaxValue)$
5. Replace zero values in education micro data by adding 0.1 rather than 1 year to all entries. ($E_M = E_o + 1$; $E_{M'} = E_o + 0.1$))
6. Consider the potential HDI, H_1^*

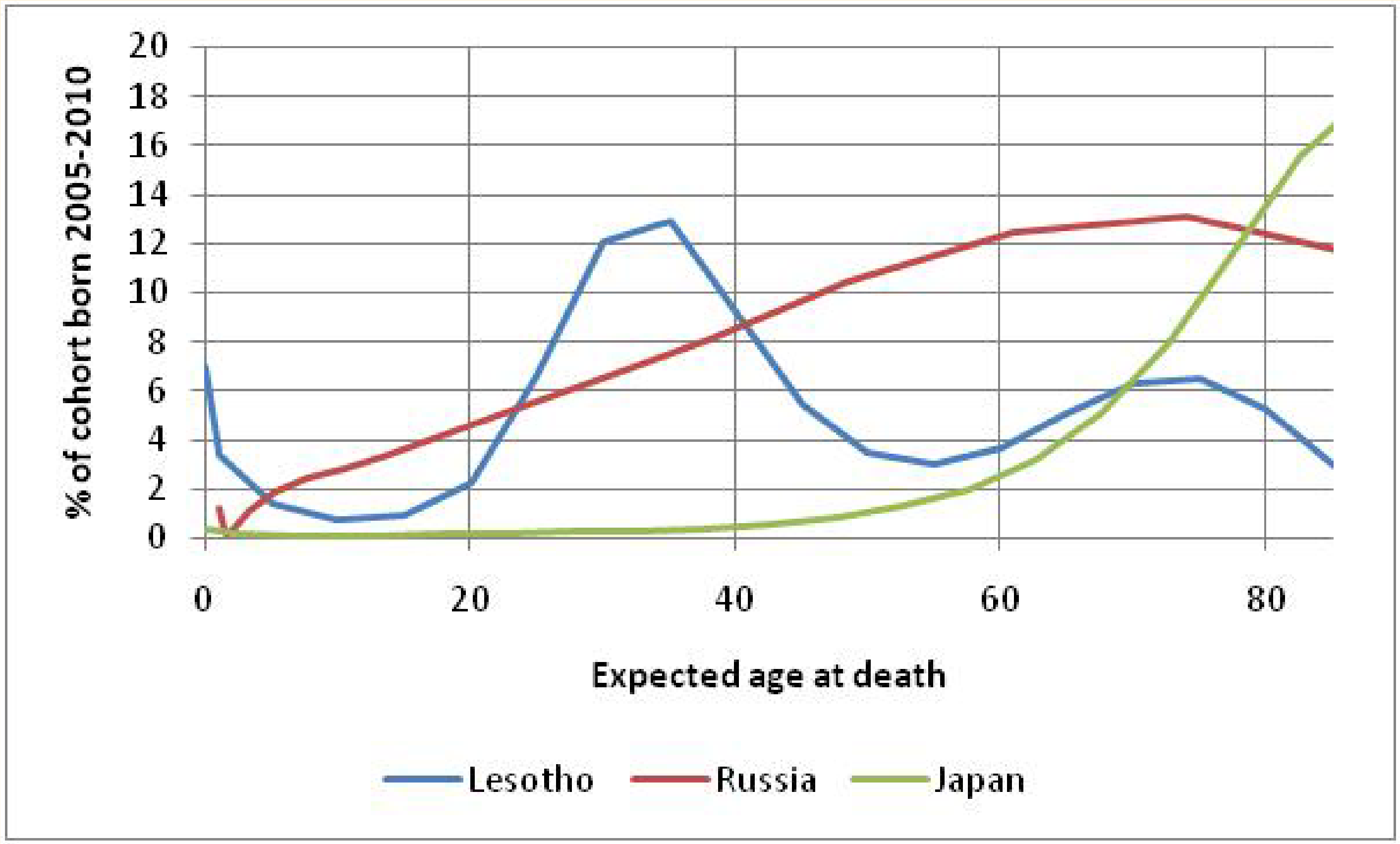
		Indicators						
		Life expectancy at birth		Mean years of schooling		Expected years of schooling		GNI per capita
		(years)		(years)		(years)		(PPP US$)
Norway		81.0		12.6		17.3		58,810
Australia		81.9		12.0		20.5		38,692
United Kingdom		79.8		9.5		15.9		35,087
Mexico		76.7		8.7		13.4		13,971
Viet Nam		74.9		5.5		10.4		2,995
Nicaragua		73.8		5.7		10.8	g	2,567
Nigeria		48.4		5.0	z	8.9	g	2,156
Haiti		61.7		4.9		6.8	g	949

Some Explorations

	IHDI	IHDI-LB-H15	IHDI-UB-Y81	IHDI-Log	Ed-geo	Ed-0-0.1	H_1*
Norway	0.901	0.921	0.764	0.980	0.907	0.973	0.965
Australia	0.826	0.910	0.718	0.965	0.833		0.899
United Kingdom	0.717	0.883	0.683	0.958	0.723	0.973	0.795
Mexico	0.435	**0.753**	0.438	0.847	0.439		0.553
Viet Nam	0.243	**0.787**	0.277	0.799	0.245		0.292
Nicaragua	0.223	0.747	0.252	0.776	0.224		0.279
Nigeria	0.111	0.329	**0.154**	0.554	0.112		**0.216**
Haiti	0.114	0.554	**0.151**	0.621	0.114		**0.171**

Problems for Future Research

1. Income Versus Consumption
2. Data Quality Health and Education
3. Inequality of Proxies vs HDI variables
4. Zero Replacements & Robustness
5. Natural Zero, Normative Cutoffs and the Ratio Scale Assumption
6. Equivalence Scales



	Japan	Lesotho	Russian Federation
Life expectancy	83.2	45.9	67.2
Atkinson measure(ε=1)	0.039	0.366	0.115