

慶應義塾大学 教養研究センター

情報の教養学講演シリーズ第3回

データで貧困を 考えよう

新保一成

慶應義塾大学商学部

2014年6月19日@日吉キャンパ

ス来往舎1Fシンポジウムス

ペース

本日の話題

1. ビッグデータと統計学について
2. 標本調査について
3. 標本調査を利用した貧困の測定について

ビッグデータ社会

- ❖ おいしいビッグデータ（朝日新聞，2014年6月8日朝刊，ザ・テクノロジー第2部AI編[下]）
 - ✓ すかいらく：「勘と経験だけでは通用しない。これからはデータサイエンスの時代」
 - ✓ スシロー：データによる予測に加えて，店長が天気などを見ながら情報を微調整。
 - ✓ プライバシーの問題，消費者が企業にコントロールされる，など未整備の問題。
- ❖ 同様の懸念は以下でも述べられている。
 - ✓ 天声人語（朝日新聞，2014年6月11日朝刊）：ビッグデータの不安。
 - ✓ 「1984年」で考えるビッグデータ社会（朝日新聞，2014年6月2日朝刊）：個人情報保護法改正との関連。
- ❖ 西垣 通，ドミニク・チェン「情報は人を自由にするか」（『現代思想』特集：ポスト・ビッグデータと統計学の時代，青土社，2014年6月号）：創造性の抑圧，標本化される社会

統計的方法によってコンピューターが 人間の思考に置き換わる？

- ❖ ビクター・マイヤー・ショーンベルガー, ケネス・クキエ 『ビッグデータの正体 情報の産業革命が世界のすべてを変える』 (講談社, 2013年)
- ❖ イアン・エアーズ 『その数学が戦略を決める』 (文藝春秋社, 2010年) : Wine tasting, Amazon, Money Ball, etc.
- ❖ 西内 啓 『統計学が最強の学問である』 (ダイヤモンド社, 2013年)
- ❖ 特集: ポスト・ビッグデータと統計学の時代 (『現代思想』, 青土社, 2014年6月号)

ビッグデータの例

- ❖ 震災ビッグデータ (NHK特集, file1 : 2013年3月3日, file2 : 9月8日, file3 : 2014年3月2日)
- ❖ スマート・グリッド
- ❖ Dominick's Database : the Kilts center at the University of Chicago
 - ✓ Dominick's Finer Foodsのweekly store-level scanner data
- ❖ 特徴 : 調査票, 標本設計を必要としない. 全データが解析の対象? 分析手法としてのデータマイニングとの違いは?

鍵

- ❖ 母集団概念を想定できるか.
- ❖ 標本は母集団からの無作為抽出と考えられるか.
 - ✓ 大数の法則：サンプルサイズが大きくなれば標本平均が母集団平均に近づく確率が1になる.
 - ✓ 中心極限定理：サンプルサイズが大きくなれば標本平均の分布は正規分布で近似できる.

標本調査

- ❖ 統計学を用いることによって部分から全体を推測する。
- ❖ 母集団分布を反映する標本を抽出できるような標本設計が最も重要であり、その上でサンプルサイズを大きくすることは望ましい。
- ❖ さらに観測誤差をできるだけ小さくすることができればさらに望ましい。

貧困 at a Glance



More developed regions 17人



India 17人



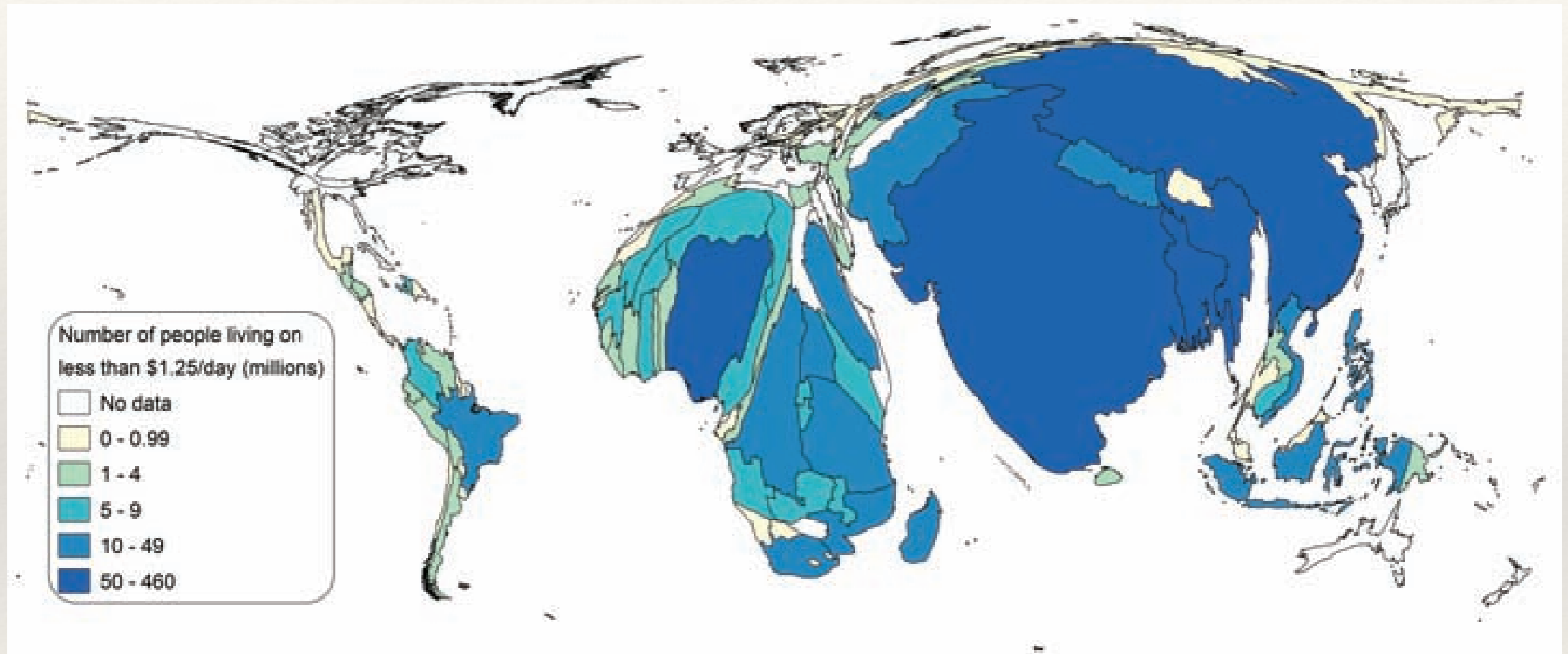
China 18人



Other less developed regions 48人

貧困 at a Glance:

Number of people living on less than \$1.25/day



Source: Atlas of World Hunger

貧困 at a Glance: WDI iPhone/iPad

The screenshot shows the World Bank's 'Poverty and Inequality Indicators' app interface on an iPad. At the top, the status bar displays 'iPad', signal strength, time '4:15', and battery level '34%'. The app header features the World Bank logo and the text 'THE WORLD BANK Working for a World Free of Poverty' on the left, and 'Poverty and Inequality Indicators' on the right. Below the header is a descriptive paragraph: 'The World Bank collection of poverty and inequality indicators is compiled from officially recognized international sources. It presents the most current and accurate data on poverty and inequality in developing countries, and includes national, regional and global estimates. To view the latest data, you need an internet connection. More on poverty at www.worldbank.org/poverty'. Below this text are four navigation buttons: 'By Country', 'By Group' (highlighted in red), 'Indicators', and 'Advanced Query'. A search bar with a magnifying glass icon and the text 'Search' is positioned above a scrollable list of categories. The visible categories in the list are: 'Middle East & North Africa (developing only)', 'South Asia', 'Sub-Saharan Africa (developing only)', 'Low income', 'Middle income', 'Lower middle income', 'Upper middle income', and 'Global (developing only)'. At the bottom of the screen is a dock with three icons: a question mark, a share icon, and a star icon labeled 'Favorites'.

貧困 at a Glance: WDI iPhone/iPad

iPad 4:15 34%

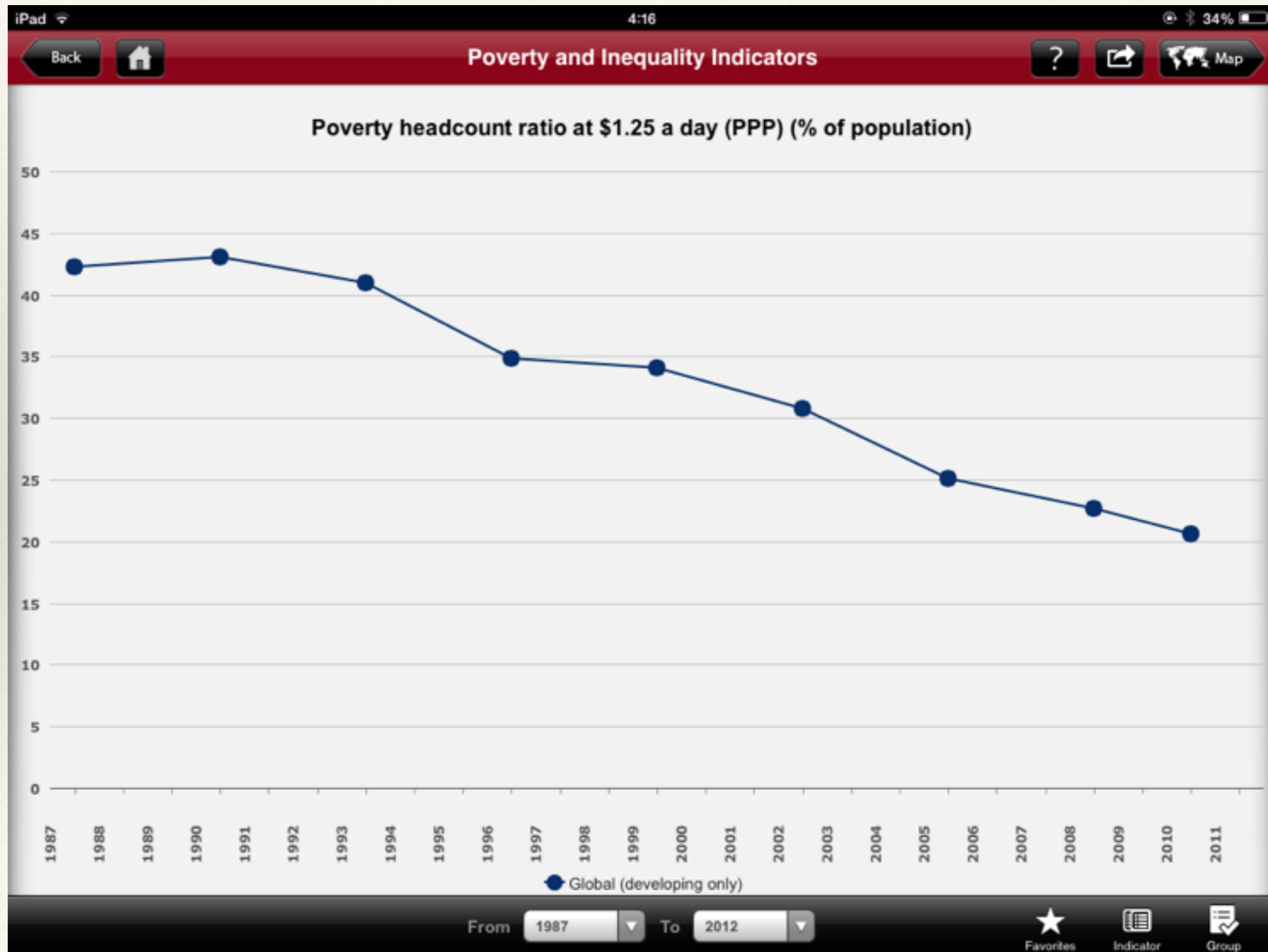
Back Home Poverty and Inequality Indicators ?

← Global (developing only) ⓘ

Indicators	2002	2005	2008	2010	MRV
\$1.25 a day (PPP) poverty line					
ⓘ Poverty headcount ratio at \$1.25 a day (PPP) (% of population)	30.8	25.1	22.7	20.6	20.6
ⓘ Poverty gap at \$1.25 a day (PPP) (%)	9.9	7.7	6.9	6.2	6.2
ⓘ Number of poor at \$1.25 a day (PPP) (millions)	1,639.2	1,389.5	1,301.9	1,215.0	1,215.0
\$2 a day (PPP) poverty line					
ⓘ Poverty headcount ratio at \$2 a day (PPP) (% of population)	53.5	46.9	43.3	40.7	40.7
ⓘ Poverty gap at \$2 a day (PPP) (%)	22.4	18.6	17.0	15.7	15.7
ⓘ Number of poor at \$2 a day (PPP) (millions)	2,848.4	2,595.8	2,484.9	2,395.2	2,395.2
\$2.50 a day (PPP) poverty line					
ⓘ Poverty headcount ratio at \$2.5 a day (PPP) (% of population)	62.5	56.6	52.5	50.0	50.0
ⓘ Poverty gap at \$2.5 a day (PPP) (%)	29.6	25.3	23.2	21.6	21.6
ⓘ Number of poor at \$2.5 a day (PPP) (millions)	3,330.2	3,133.3	3,017.0	2,942.3	2,942.3
\$4 a day (PPP) poverty line					
ⓘ Poverty headcount ratio at \$4 a day (PPP) (% of population)	77.5	72.9	68.6	66.4	66.4
ⓘ Poverty gap at \$4 a day (PPP) (%)	45.1	40.5	37.6	35.7	35.7

Favorites Group Compare

貧困 at a Glance: WDI iPhone/iPad



貧困の測定

- ❖ あるスーパーマーケットに来店する顧客の情報だけでは全体を把握することはできない。
- ❖ 標本調査の必要性。
 - ✓ 国の統計機関や世界銀行などの国際機関が実施する『家計調査』(Household Survey)や生計費調査(Living Standard Survey)

Chen & Ravallion(2010)の分析

- ❖ 1979年から2006年の115カ国に関する675の家計調査を駆使。2005年の推定値には、123万家計への調査が利用されている。
- ❖ 国連ICP(International Comparison Project)と世界銀行によるPPP(Purchasing Power Parity, 購買力平価)によって貧困線を同定。
- ❖ Shaohua Chen and Martin Ravallion(2010) “The Developing World is Poorer than We Thought, But No Less Successful in the Fight against Poverty,” *Quarterly Journal of Economics*.

Chen & Ravallion(2010)の分析

TABLE III

REGIONAL BREAKDOWN OF HEADCOUNT INDEX FOR INTERNATIONAL POVERTY LINES OF \$1.00–\$2.50 A DAY OVER 1981–2005

Region	1981	1984	1987	1990	1993	1996	1999	2002	2005
	(a) % living below \$1.00 a day								
East Asia and Pacific	66.8	49.9	38.9	39.1	35.4	23.4	23.5	17.8	9.3
Of which China	73.5	52.9	38.0	44.0	37.7	23.7	24.1	19.1	8.1
Eastern Europe and Central Asia	0.7	0.6	0.5	0.9	2.1	2.5	3.1	2.7	2.2
Latin America and Caribbean	7.7	9.2	8.9	6.6	6.0	7.3	7.4	7.7	5.6
Middle East and North Africa	3.3	2.4	2.3	1.7	1.5	1.6	1.7	1.4	1.6
South Asia	41.9	38.0	36.6	34.0	29.3	29.1	26.9	26.5	23.7
Of which India	42.1	37.6	35.7	33.3	31.1	28.6	27.0	26.3	24.3
Sub-Saharan Africa	42.6	45.2	44.1	47.5	46.4	47.6	47.0	43.8	39.9
Total	41.4	34.4	29.8	29.5	27.0	23.1	22.8	20.3	16.1
	(b) % living below \$1.25 a day								
East Asia and Pacific	77.7	65.5	54.2	54.7	50.8	36.0	35.5	27.6	16.8
Of which China	84.0	69.4	54.0	60.2	53.7	36.4	35.6	28.4	15.9
Eastern Europe and Central Asia	1.7	1.3	1.1	2.0	4.3	4.6	5.1	4.6	3.7
Latin America and Caribbean	11.5	13.4	12.6	9.8	9.1	10.8	10.8	11.0	8.2
Middle East and North Africa	7.9	6.1	5.7	4.3	4.1	4.1	4.2	3.6	3.6
South Asia	59.4	55.6	54.2	51.7	46.9	47.1	44.1	43.8	40.3
Of which India	59.8	55.5	53.6	51.3	49.4	46.6	44.8	43.9	41.6
Sub-Saharan Africa	53.7	56.2	54.8	57.9	57.1	58.7	58.2	55.1	50.9
Total	51.8	46.6	41.8	41.6	39.1	34.4	33.7	30.6	25.2

Source: Chen & Ravallion(2010)

Chen & Ravallion(2010)の分析

TABLE III
(CONTINUED)

Region	1981	1984	1987	1990	1993	1996	1999	2002	2005
	(c) % living below \$2.00 a day								
East Asia and Pacific	92.6	88.5	81.6	79.8	75.8	64.1	61.8	51.9	38.7
Of which China	97.8	92.9	83.7	84.6	78.6	65.1	61.4	51.2	36.3
Eastern Europe and Central Asia	8.3	6.5	5.6	6.9	10.3	11.9	14.3	12.0	8.9
Latin America and Caribbean	22.5	25.3	23.3	19.7	19.3	21.8	21.4	21.7	16.6
Middle East and North Africa	26.7	23.1	22.7	19.7	19.8	20.2	19.0	17.6	16.9
South Asia	86.5	84.8	83.9	82.7	79.7	79.9	77.2	77.1	73.9
Of which India	86.6	84.8	83.8	82.6	81.7	79.8	78.4	77.5	75.6
Sub-Saharan Africa	74.0	75.7	74.2	76.2	76.0	77.9	77.6	75.6	73.0
Total	69.2	67.4	64.2	63.2	61.5	58.2	57.1	53.3	47.0
	(d) % living below \$2.50 a day								
East Asia and Pacific	95.4	93.5	89.7	87.3	83.7	74.9	71.7	62.6	50.7
Of which China	99.4	97.4	92.4	91.6	86.5	76.4	71.7	61.6	49.5
Eastern Europe and Central Asia	15.2	12.5	11.2	12.0	15.1	18.3	21.4	17.8	12.9
Latin America and Caribbean	29.2	32.4	29.6	26.0	25.9	28.8	28.0	28.4	22.1
Middle East and North Africa	39.0	34.8	34.6	31.2	31.4	32.5	30.8	29.5	28.4
South Asia	92.6	91.5	90.8	90.3	88.6	88.5	86.7	86.5	84.4
Of which India	92.5	91.5	90.8	90.2	89.9	88.7	87.6	86.9	85.7
Sub-Saharan Africa	81.0	82.3	81.0	82.5	82.5	84.2	83.8	82.5	80.5
Total	74.6	73.7	71.6	70.4	69.2	67.2	65.9	62.4	56.6

Source: Chen & Ravallion(2010)

補論: PPP(Bic Macレート)

国	Bic Mac 価格	PPP 購買力平価	為替 レート (\$1あたり)
USA	\$4.62		
UK	£2.79	0.60	0.59
JPN	¥310	67.10	102.45
CHN	元16.6	3.59	6.18
IND	95Rupees	20.56	60.99

<http://www.economist.com/content/big-mac-index>

標本調査の実際：日本の家計調査

- ❖ 家計調査は統計局が統計法に基づいて実施している。
- ❖ 層化3段抽出法によって全国からおよそ9000世帯を無作為抽出する。
 - ✓ 第1層：直近の国勢調査により全国の市区町村を同じ性質になるようにグループ分けし、各グループから1市区町村を選ぶ。
 - ✓ 第2層：各調査市区町村から調査地区を選ぶ。これを単位区という。
 - ✓ 第3層：各単位区から2人以上世帯を6世帯、単身世帯を2世帯、乱数表によって無作為抽出する。
- ❖ 2人以上世帯は6カ月、単身世帯は3カ月家計簿を記入する。
- ❖ 毎月、2人以上世帯は1/6ずつ、単身世帯は1/3ずつ交替する。

標本調査の実際：日本の家計調査

記入例 2

支出の記入のしかた

数量・単位の記入

*購入した品目の数量は、1山、1袋、1尾、1本などと記入せず400グラム、1800ミリリットル（又は、400g、1800ml）などのように量目で記入します。

各期の第1日目（1日又は16日）に、前日から繰り越した手持ち金額を記入します。

/ 日(火曜日)

品名などの書き方

*「うどん・そば」は、ゆでたものか干したものかなどを区別して記入します。

*「魚」「肉」「野菜」「パン」などではなく、品名を具体的に記入します。

*誰が使うものかを記入します。

*何に使うためかを記入します。

*月ぎめの牛乳、新聞などは、代金を支払った日に「現金支出」欄に記入します。
(日々の掛買いと はしません。)

I 現金収入又は現金支出

(1) 収入の種類又は支出の品名及び用途	(2) 現金収入(円)	(3) 数量 単位	(4) 現金支出(円)
1 ゆでうどん		400 匁	320
2 あじ(生)		430 匁	330
3 かき(貝)		460 匁	400
4 豚肉		330 匁	630
5 ほうれん草		300 匁	186
6 バターロール(8コ入り)		280 匁	200
7 靴下(世帯主)		2 足	1,050
8 ポロシャツ(長女)		1 枚	2,625
9 リンゴ(病気見舞い)		1,950 匁	1,800
10 可し出前(来客用)		4 人前	4,800
11 エアコン月賦支払初回分			26,000
12 酒屋掛買ひ支払10回分			4,500
13 牛乳代10月分(200ml30本)		6,000 ml	2,835
14 〇〇新聞 10月分			3,925
15			
合計			49,601

前期からの繰越金(手持ち現金) 23,060 円

本日の現金残高 33,459 円

標本調査の実際：インド家計調査

- ❖ NSSO(National Sample Survey Organization)が実施する。
- ❖ 5年毎に層化3段抽出法で都市4万世帯，農村8万世帯を訪問調査。食料，エネルギー，医療・教育などのサービスについては過去30日，衣服，耐久消費財については過去365日の消費支出を面談調査。中間年のサンプルサイズは半分。

標本調査の実際：インド家計調査

[13] perception of household regarding sufficiency of food				
1. do all members of your household 'get two square meals every day'? (yes: every month of the year-1, some months of the year -2; no: no month of the year-3)				
2. if code is 2 in item 1, during which calendar months did any member of the household not 'get two square meals every day'? (applicable month codes may be recorded in the box spaces: Jan-01, Feb-02, Mar-03, Apr-04, May-05, Jun-06, Jul-07, Aug-08, Sep-09, Oct-10, Nov-11, Dec-12)				
3. whether information on item 1 was actually obtained from the informant (yes-1, no-2)				

標本調査の実際：インド家計調査

[3] household characteristics									
1. household size					during July 08 to June 09	14. land cultivated (0.000 ha)			
2. principal industry (NIC-2004)	description:					15. land irrigated (0.000 ha)			
	code (5-digit)								
3. principal occupation (NCO-2004)	description:				primary source of energy for	16. cooking (code)			
	code (3-digit)					17. lighting (code)			
4. household type (code)					18. dwelling unit code (<i>owned-1, hired-2, no dwelling unit-3, others-9</i>)				
5. religion (code)									
6. social group (code)					19. is any member of the household a regular salary earner? (<i>yes-1, no -2</i>)				
7. whether owns any land (<i>yes-1, no -2</i>)									
8. if yes in item 7 , type of land owned (<i>homestead only – 1, homestead and other land – 2, other land only – 3</i>)					20. did the household perform any ceremony during the last 30 days? (<i>yes – 1, no – 2</i>)				
land as on the date of survey (in 0.000 hectares)									
9. owned					21. no. of meals served to non-household members during the last 30 days				
10. leased-in									
11. otherwise possessed (neither owned nor leased-in)					22. whether the household has access to internet at home on the date of survey (<i>yes-1, no-2</i>)				
12. leased-out									
13. total possessed [items (9+10+11-12)]									

Source: Schedule 1.0, National Sample Survey Organization, Government of India

標本調査の実際：インド家計調査

[5.1] consumption of cereals, pulses, milk and milk products, sugar and salt during the last 30 days ended on						
item	code	consumption out of home produce		total consumption		source code
		quantity@ (0.000)	value (Rs.)	quantity@ (0.000)	value (Rs.)	
(1)	(2)	(3)	(4)	(5)	(6)	(7)
rice – PDS	101					1
rice – other sources	102					
chira	103					
khoi, lawa	104					
muri	105					*
other rice products	106					*
wheat/ atta – PDS	107					1
wheat/ atta – other sources	108					
maida	110					
suji, rawa	111					*
sewai, noodles	112					*
bread (bakery)	113					*
other wheat products	114					*

Source: Schedule 1.0, National Sample Survey Organization, Government of India

インドの識字率

- ❖ 農村では、女性の55%、男性の36%が文盲。
- ❖ 都市でも、女性の31%、男性の20%が文盲。
- ❖ ただし、プライマリに入学する子供の数が増えているので、15歳から24歳にかぎると、女性の25%、男性の12%が文盲。

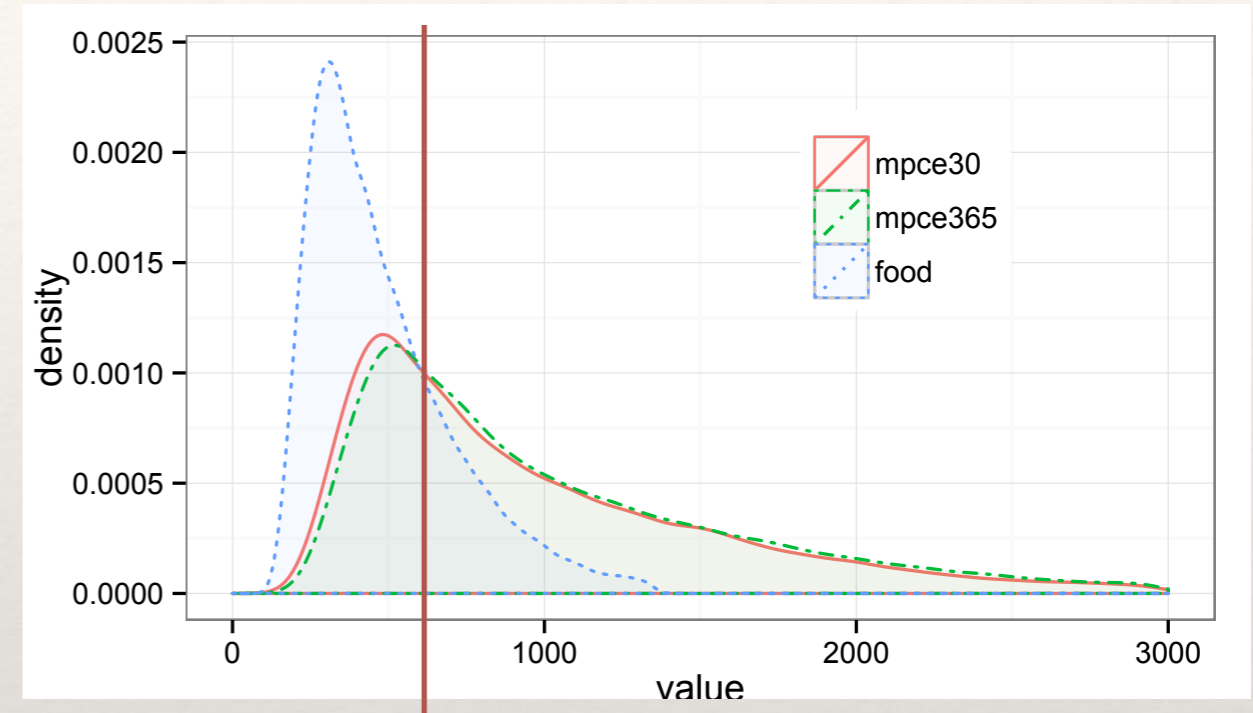
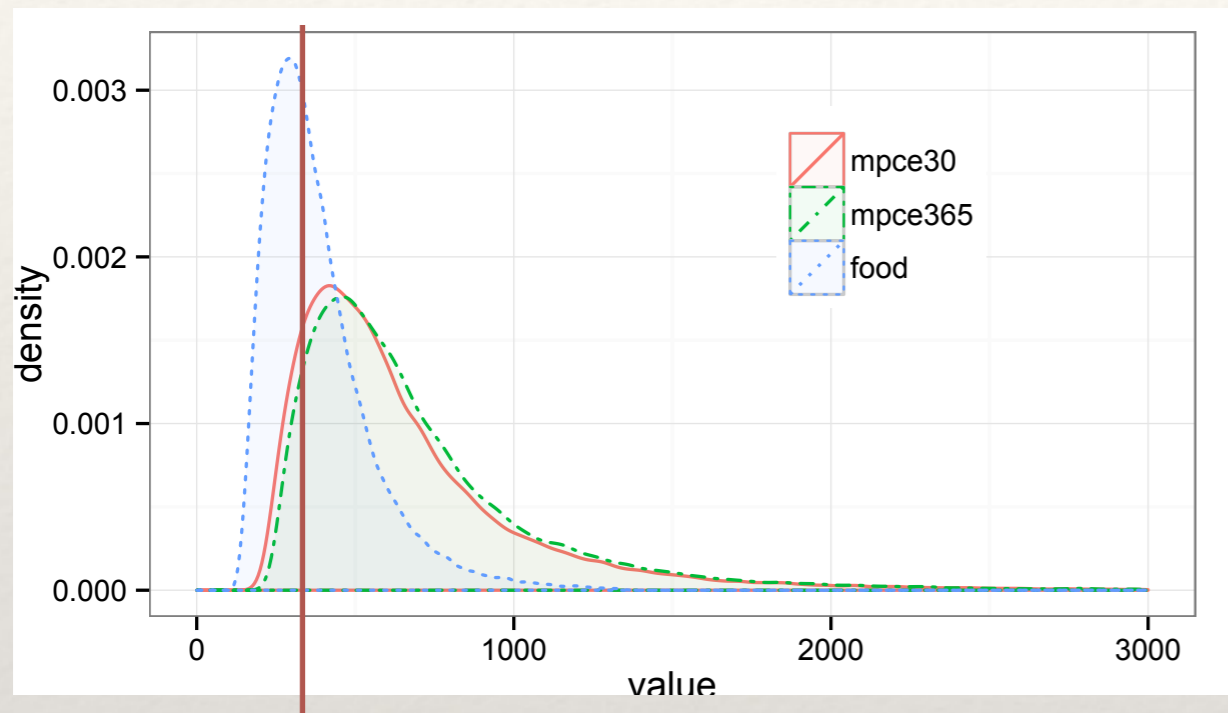


Photo by Kazushige Shimpo, 2011 新保研India Study Tour

標本調査の実際：インド家計調査



新保研 2012 India Study Tourより



	Rural				Urban			
	平均値	中央値	最小値	最大値	平均値	中央値	最小値	最大値
mpce30	664.1	553.1	200.2	3979.0	1000.4	793.8	207.6	3998.0
mpce365	706.1	590.9	250.0	8849.0	1080.1	845.4	250.2	8881.7
food	382.3	343.6	135.0	1348.0	479.9	421.9	135.0	1350.0

貧困に対する脆弱性の概念

- ❖ ある時点で貧困でない者が将来貧困に陥ってしまう確率, あるいは, ある時点で貧困である者が将来も貧困に留まってしまう確率.
- ❖ 前者には貧困防止政策, 後者には貧困緩和政策が必要.

貧困に対する脆弱性の要因

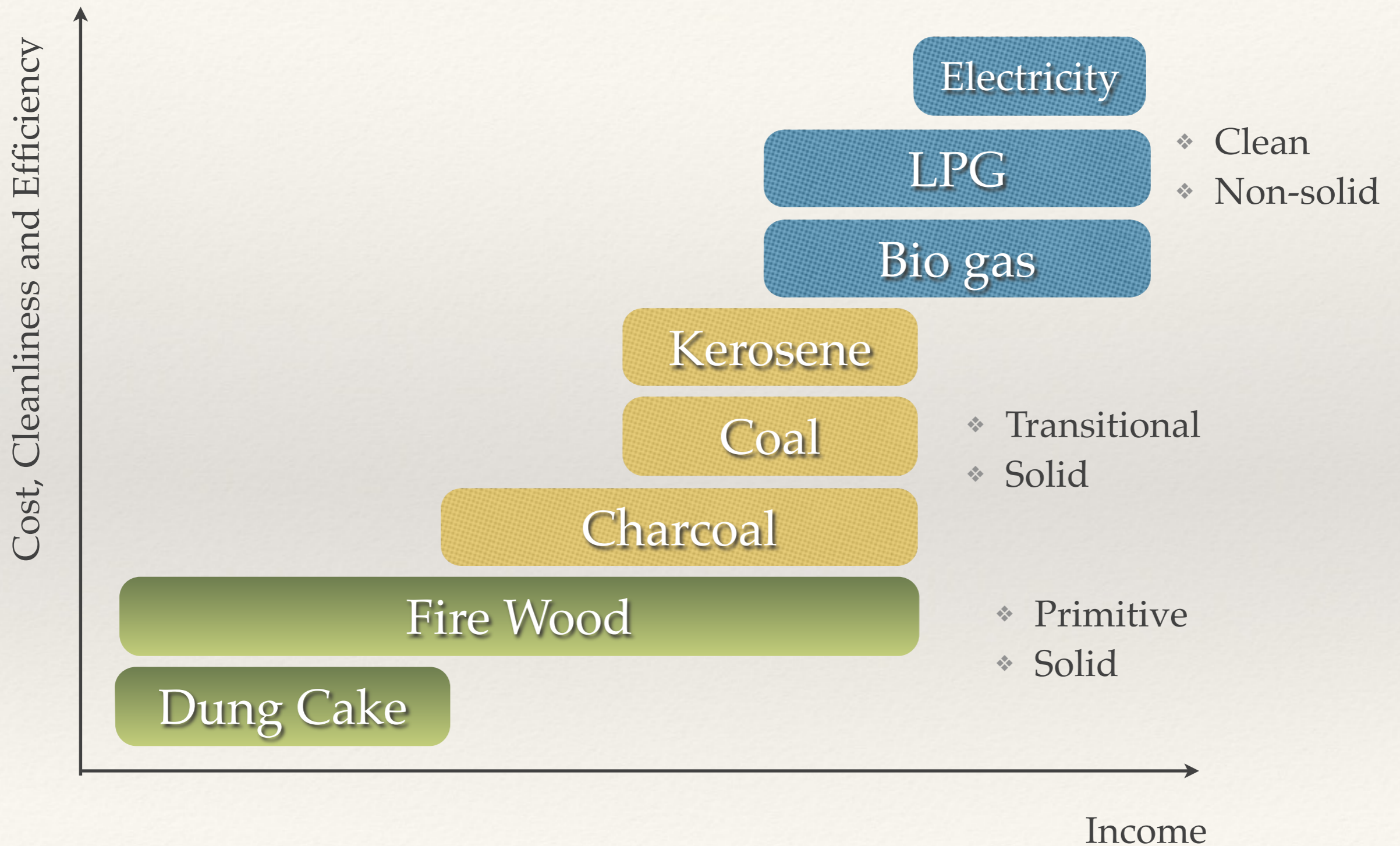
- ❖ 人的資本のレベルが低いこと，また情報へのアクセスが制限されていること。
- ❖ 身体的・精神的な不自由に苦しんでいること．生産的な資産を持っていないこと。
- ❖ 社会的に排除されている，もしくは適切な社会的ネットワークを構築できていないこと。
- ❖ 信用市場や保険市場へのアクセスが制限されていること。
- ❖ 農業環境の悪い地域，あるいは，自然資源が希少な地域に暮らしていること。
- ❖ 労働需要の少ないコミュニティーに属していること。

貧困に陥る例

- ❖ 牛乳をよく出す雌牛を飼っていて、賃労働収入を牛乳販売で補ってきた土地なし農業労働者がいる。あるとき交通事故に遭い、その年の賃労働収入がまったく得られなくなってしまう。彼は、日々の暮らしのためにその雌牛を売却して急をしのいだ。しかし、ケガが治癒しても、その後の所得源は賃労働収入のみになってしまう。このときに、信用市場や保険市場へのアクセスが可能であったならば、その者は大切な資産であった雌牛を売却せずに済んだかもしれない。

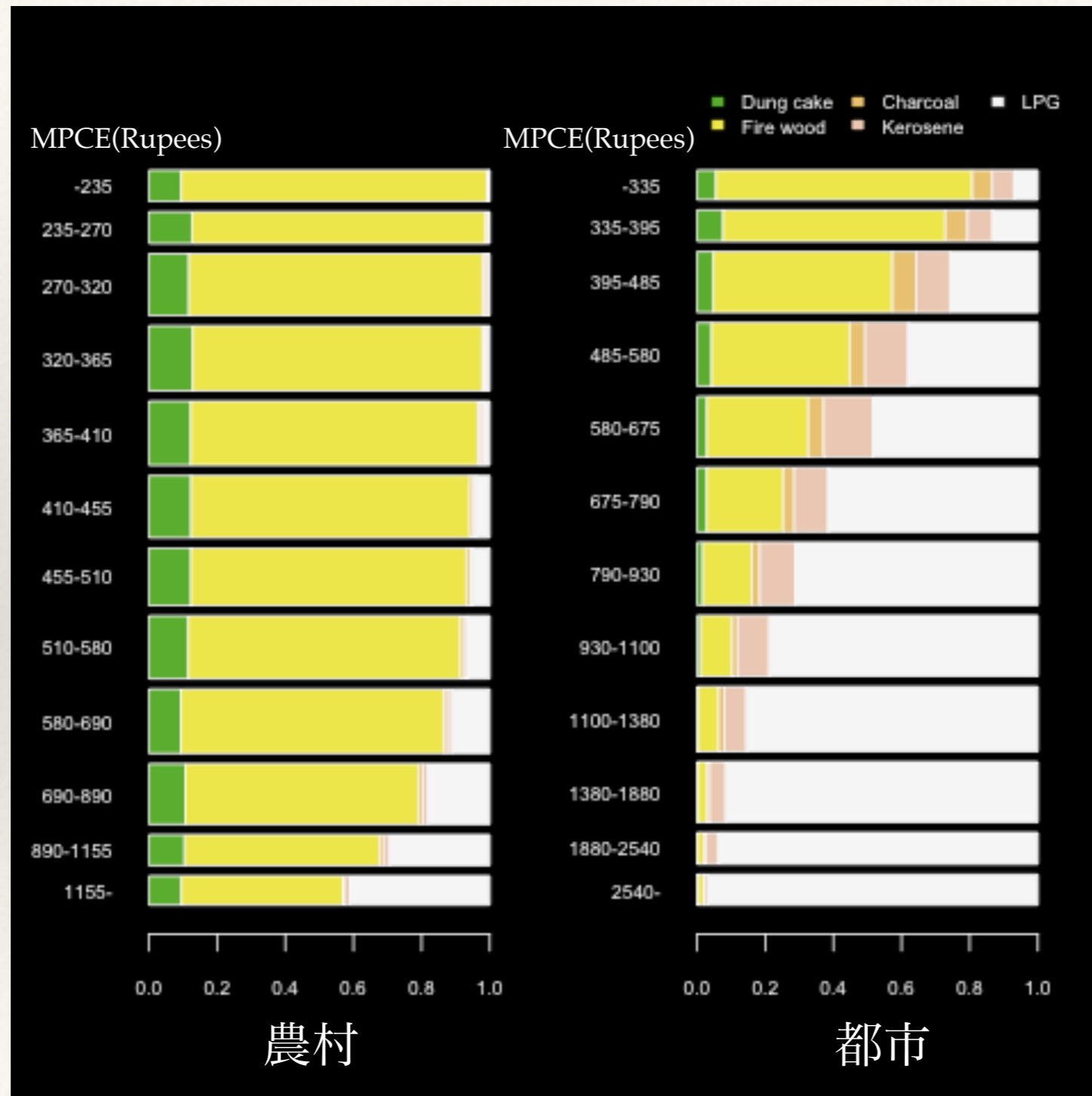
脆弱性の要因

Energy Ladder of Cooking Stove



脆弱性の要因

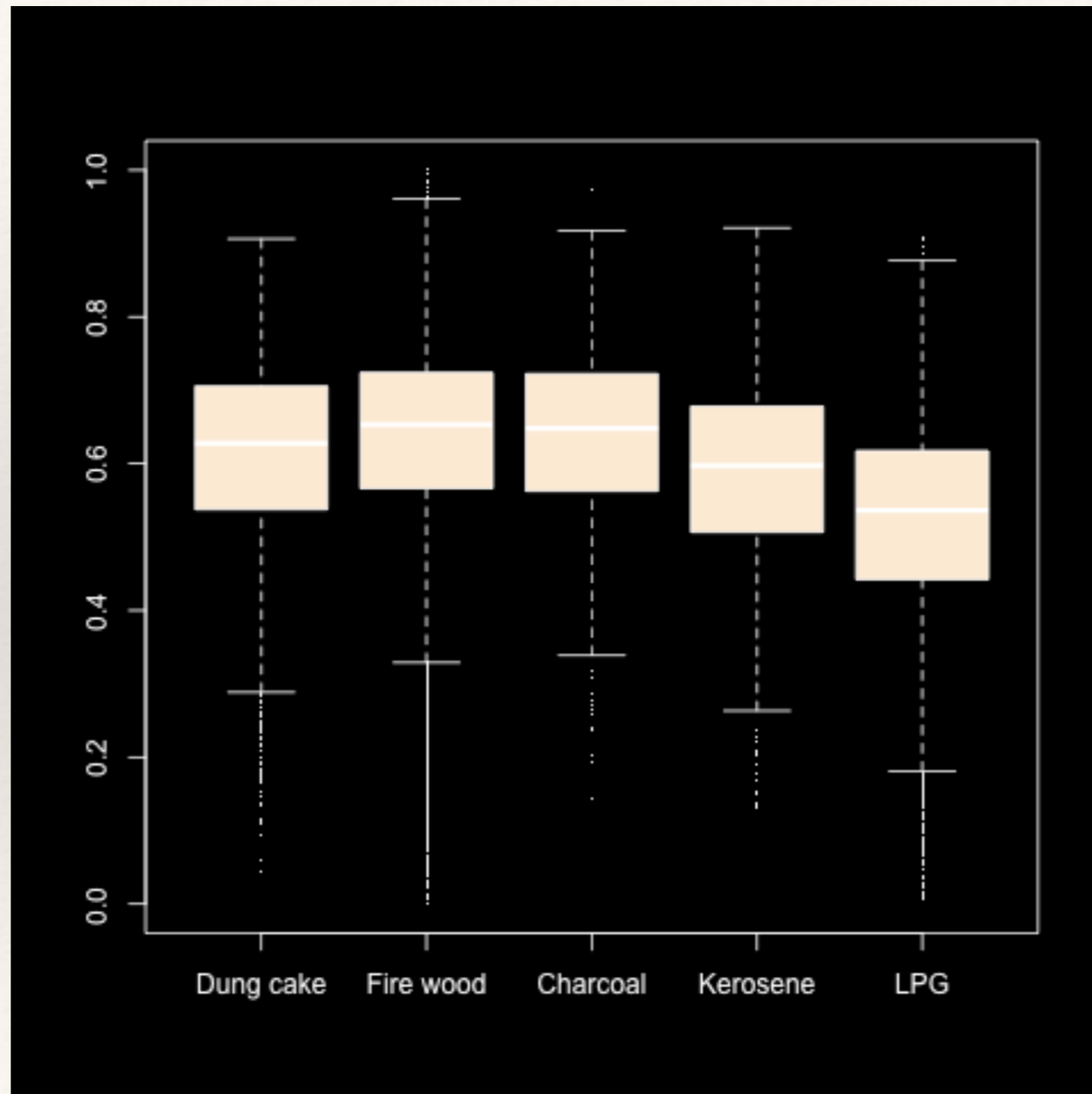
MPCE別調理用燃料



NSS 61stから作成, 2004-2005年

脆弱性の要因

調理用燃料別エンゲル係数（農村）



NSS 61stから作成, 2004-2005年

脆弱性の要因

教育の機会

日本の学校制度			インドの学校制度		
小学校	6	義務教育	プライマリ	6	ほぼ全員
中学校	12		ミドル	11	プライマリを終了するのは半数程
高等学校	15	95%	セカンダリ	14	70%
大学	18	50%	大学	18	23%

脆弱性の要因

農村の女性・女子には、学習・就労する機会に乏しい

- ❖ 農村では、電子レンジで「チン」して食べられるような便利な食べ物を口にできません。なんてたって、電気がないんですから・・・
- ❖ 朝，昼，晩の3度，材料から時間をかけて調理しなければなりません。
- ❖ 調理するための燃料，「薪」や「牛糞」を毎日毎日集めなければなりません。
- ❖ そのために1日に8時間を使います。だから，勉強する時間，働く時間を十分に取ることができません。
- ❖ それだけではなく，伝統的な厨房設備を使うことによる健康被害（室内公害）も深刻です。



Photo by Kazushige Shimpo, 2010



Photo by Kazushige Shimpo, 2010

脆弱性の要因

電力にアクセスできない人口

地域	農村	都市	全国
アフリカ	466	121	587
サハラ以南アフリカ	465	120	585
開発途上アジア	716	82	798
中国	8	0	8
インド	380	23	403
その他	328	59	387
ラテンアメリカ	27	4	31
開発途上国	1229	210	1439
世界	1232	210	1442



Photo by Kazushige Shimpo, 2010



Photo by Kazushige Shimpo, 2010



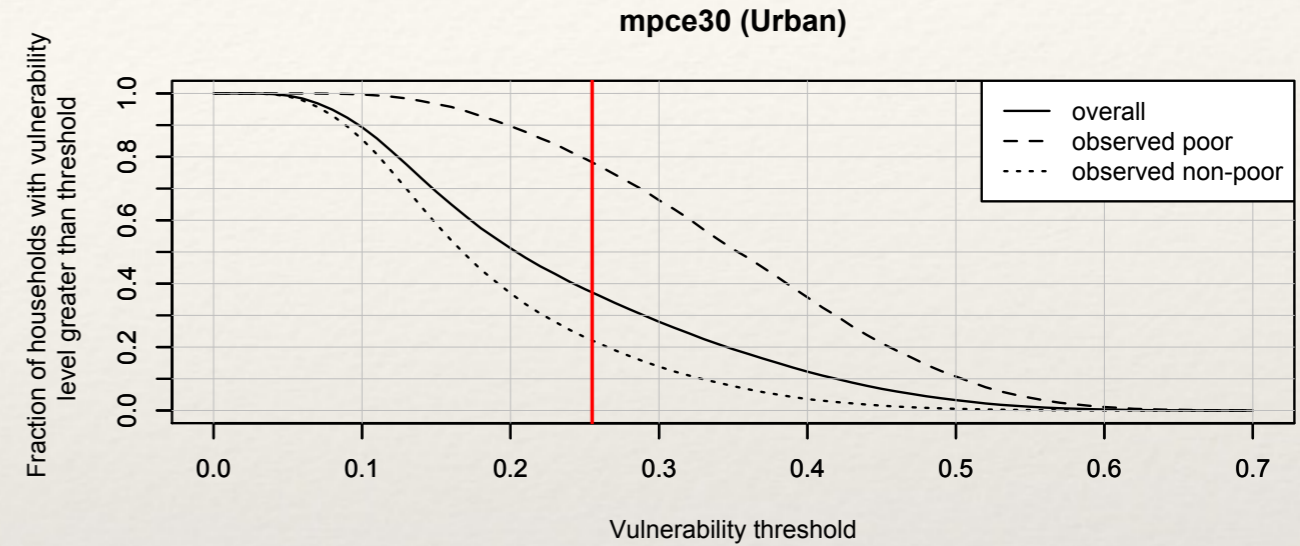
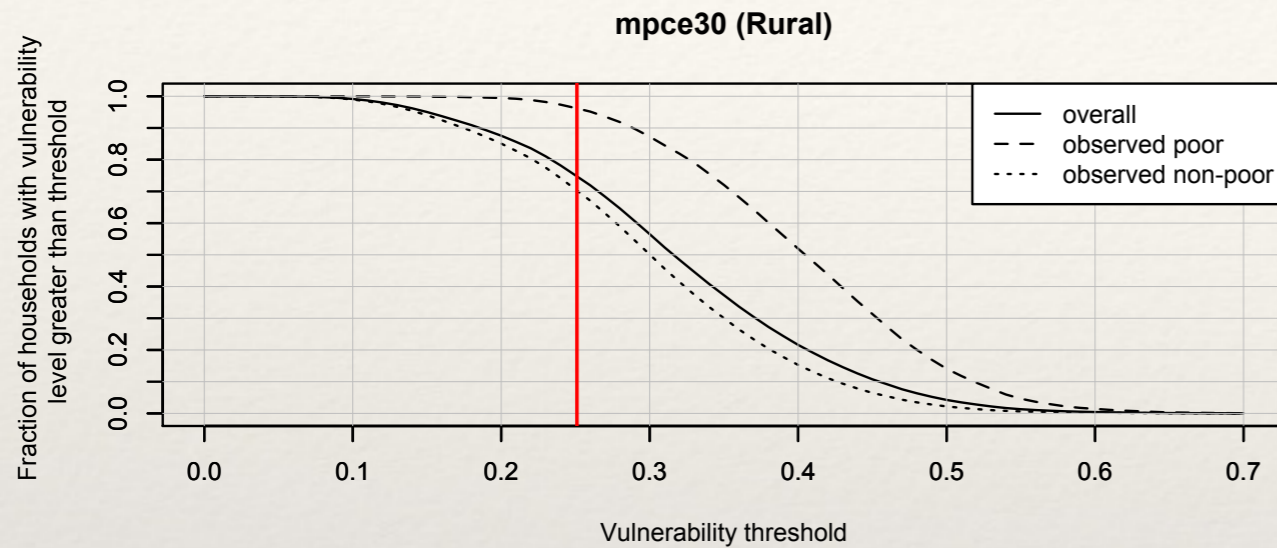
Photo by Kazushige Shimpo, 2010

2013 新保研India Study Tour

灯りがあれば
夜に働くことが
できる



貧困に対する脆弱性の測定結果



	Rural			Urban		
	overall	observed poor	observed non-poor	overall	observed poor	observed non-poor
mpce30						
貧困率	0.251	1.000	0.000	0.255	1.000	0.000
平均脆弱性指標	0.319	0.405	0.302	0.234	0.353	0.190
脆弱家計の割合	0.747	0.962	0.702	0.373	0.783	0.221

“Thank you for listening.”

–Kazushige Shimpo