Angus Deaton, Princeton University
Nobel Symposium, September 3rd, 2012

WELLBEING: MEASUREMENT & CONCEPTS

Wellbeing

- All of the things that make life worth living
  - Income (consumption, wealth, material WB)
  - Health (mortality, morbidity)
  - Education
  - Democratic participation
  - Psychological experience: depression, enjoyment
- Complete orderings
  - Utility functions, Human Development Index, Happiness, Money metric utility
- Partial orderings
  - Sen: we can say a lot without complete orderings
  - Focus here, because most concern is with measurement of components, with a view to how they fit together

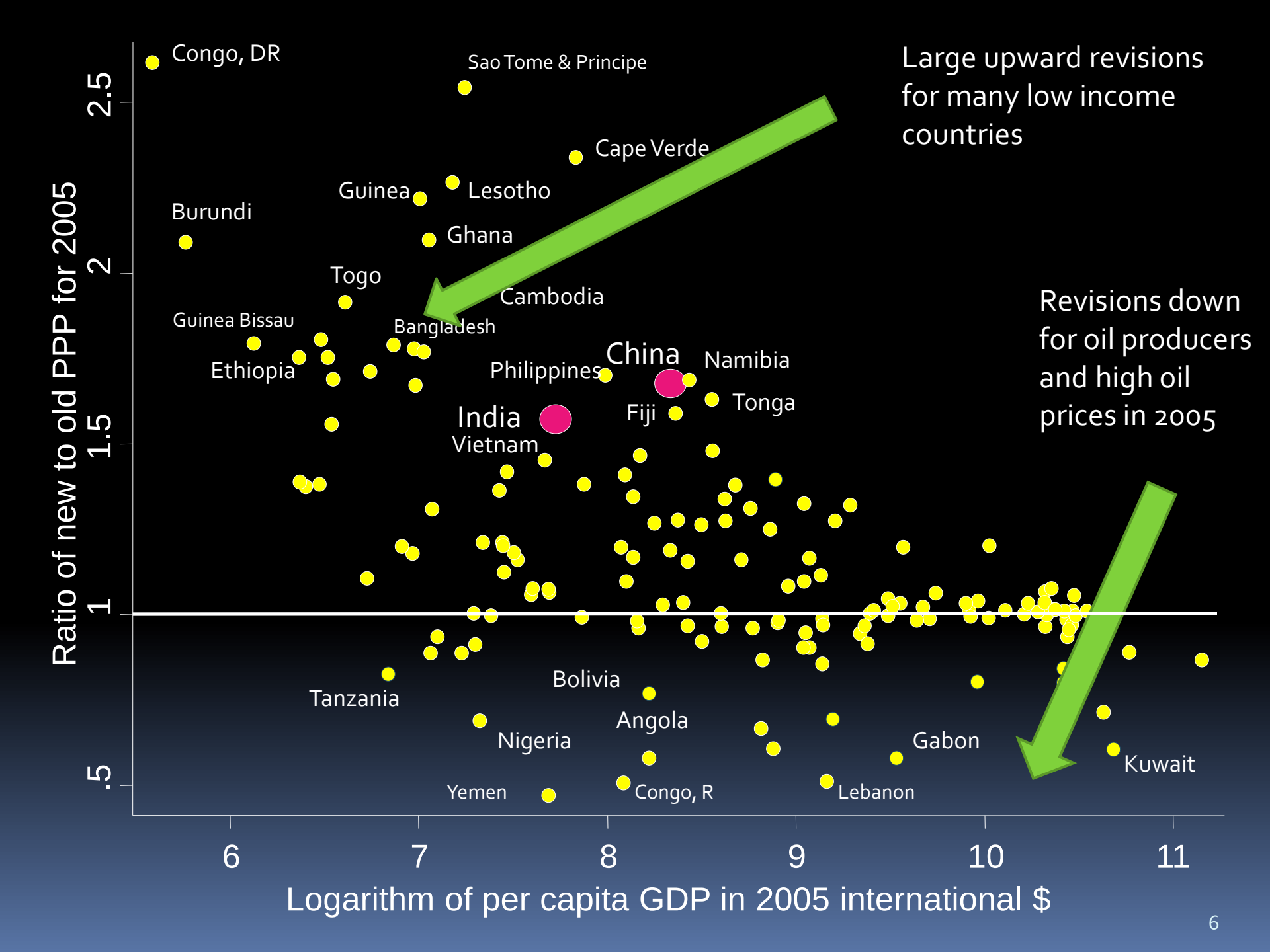
This talk

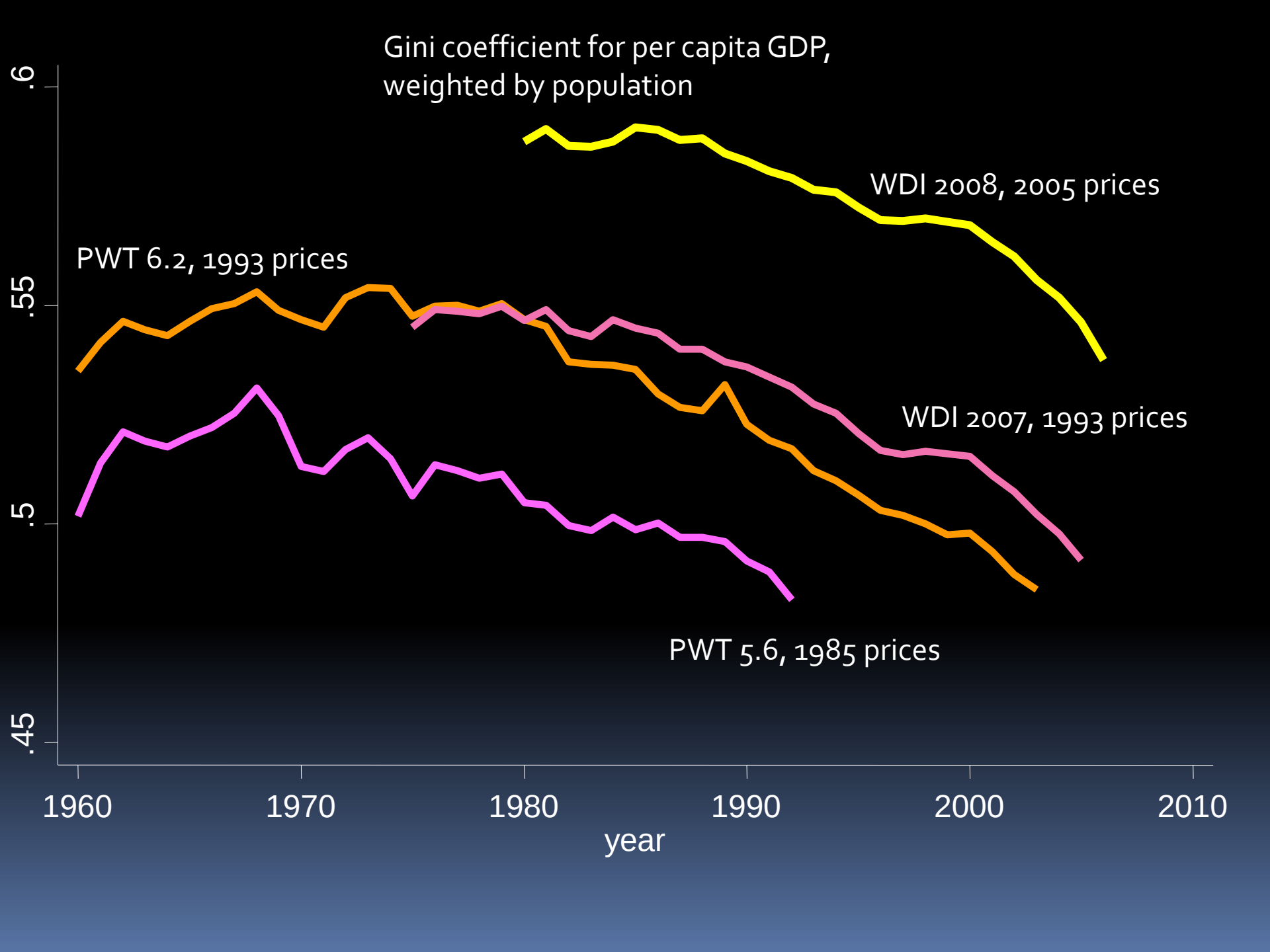
- Mostly measurement
- Three areas
  - Material well-being: purchasing power price-adjusted incomes that underpin almost all measures
    - Understanding PPPs
    - Implications for global poverty & inequality
    - Outstanding puzzles
  - Health (mortality and life expectancy, fertility, and a little about morbidity)
  - Happiness
    - Does it provide a useful complete ordering?
    - Can it be measured ?
    - What good is happiness?

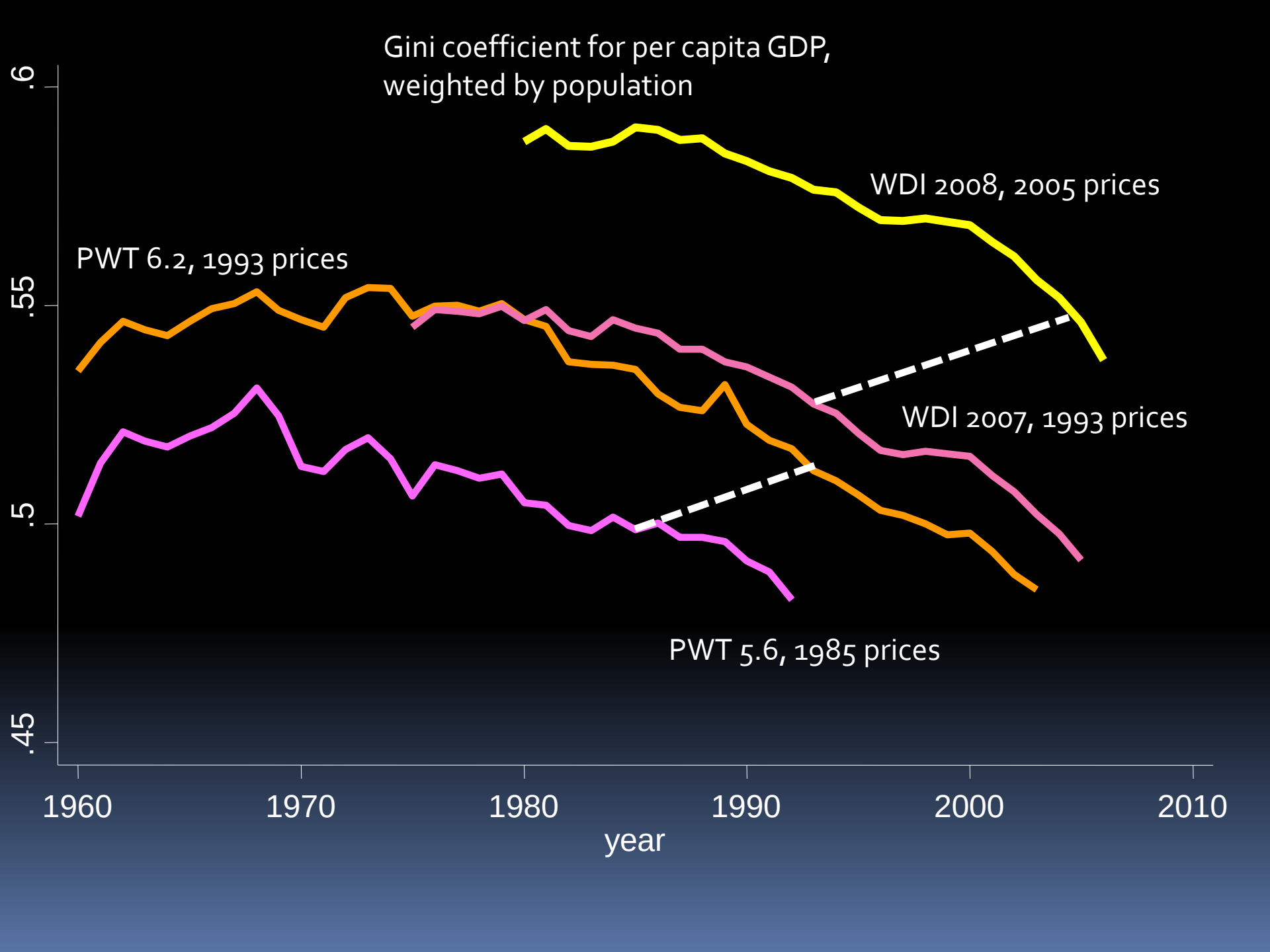
PRICE INDEXES AND GLOBAL INCOME COMPARISONS

PPP exchange rates

- The Penn World Table is the basis for the income comparisons in almost all global calculations
  - Or other numbers based on ICP price collection
- The ICP collects prices of individual goods around the world
  - Rounds in 1993-1995: PWT 6.X
  - 2005: PWT 7.0
  - 2011 (not yet released)
- Used to construct multilateral price indexes
  - Which deflate local estimates of income, consumption, etc. measured in local currencies
- Between rounds, and waiting for rounds, e.g. post 2005
  - Local CPIs (or implicit price deflators) are used to update
  - PPP between, say, US and India, should track relative growth rates of CPIs in US and India


Large upward revisions for many low income countries
Revisions down for oil producers and high oil prices in 2005
Congo, DR
Sao Tome & Principe
Cape Verde
Guinea
Lesotho
Burundi
Ghana
Togo
Cambodia
Guinea Bissau
Bangladesh
Ethiopia
Philippines
China
Namibia
Tonga
India
Fiji
Vietnam
Tanzania
Bolivia
Angola
Nigeria
Gabon
Kuwait
Yemen
Congo, R
Lebanon
Ratio of new to old PPP for 2005
.5
1
1.5
2
2.5
6
7
8
9
10
11
Logarithm of per capita GDP in 2005 international $


Gini coefficient for per capita GDP, weighted by population
WDI 2008, 2005 prices
PWT 6.2, 1993 prices
WDI 2007, 1993 prices
PWT 5.6, 1985 prices
.6
.55
.5
.45
1960
1970
1980
1990
2000
2010
year


Gini coefficient for per capita GDP,
weighted by population
WDI 2008, 2005 prices
PWT 6.2, 1993 prices
WDI 2007, 1993 prices
PWT 5.6, 1985 prices
.6
.55
.5
.45
1960
1970
1980
1990
2000
2010
year

Why?

- We do not understand, yet it is crucial for understanding changes in global income inequality
  - Not important for global poverty, because $1 a day line is average of poor country lines & depends only on PPP rates *among* poor countries, which are not much affect
    - World Bank switch from $1 to $1.25 is an increase in the line, and not consequence of PPP revision
- Note this is *not* Balassa-Samuelson
  - BS says ER and CPI change at different rates with growth
  - Because of non-traded goods, and rising productivity in traded-goods sector
  - But PPPs are updated by CPIs, not exchange rates
- Number of technical, tedious, but important details
  - E.g. treatment of trade-balance differs in ICP and NIPA
  - Which is why oil exporters got to be relatively richer in 2005

$\ln P_t^C$

National and international

- Suppose that everyone uses Törnqvist type indexes

$$\ln P_t^C = \sum_n w_{tn}^C \ln p_{tn}^C$$

- Ignore changes in the weights (whose effects are second order) then differential CPI change is

$$d \ln P_t^B - d \ln P_t^A = \sum_n w_{tn}^B d \ln p_{tn}^B - \sum_n w_{tn}^A d \ln p_{tn}^A$$

- There is also a PPP index for B relative to A, and suppose that it , too, has the same form (isn't usually the case)

$$d \ln \pi_t^{BA} = \sum_n w_{tn}^{BA} (d \ln p_{tn}^B - d \ln p_{tn}^A)$$

- For a Törnqvist index, the weights are the average of the country weights

Aggregation bias

- The difference between the updating formula and the benchmark change is then

$$d \ln \pi_t^{BA} - (d \ln P_t^B - d \ln P_t^A) = -0.5 \sum_n (w_{tn}^B - w_{tn}^A)(d \ln p_{tn}^B - d \ln p_{tn}^A)$$

- This will be zero if the shares are the same in B and A, or if the changes in relative prices are the same in B and A, or if the changes in relative prices are orthogonal to the shares

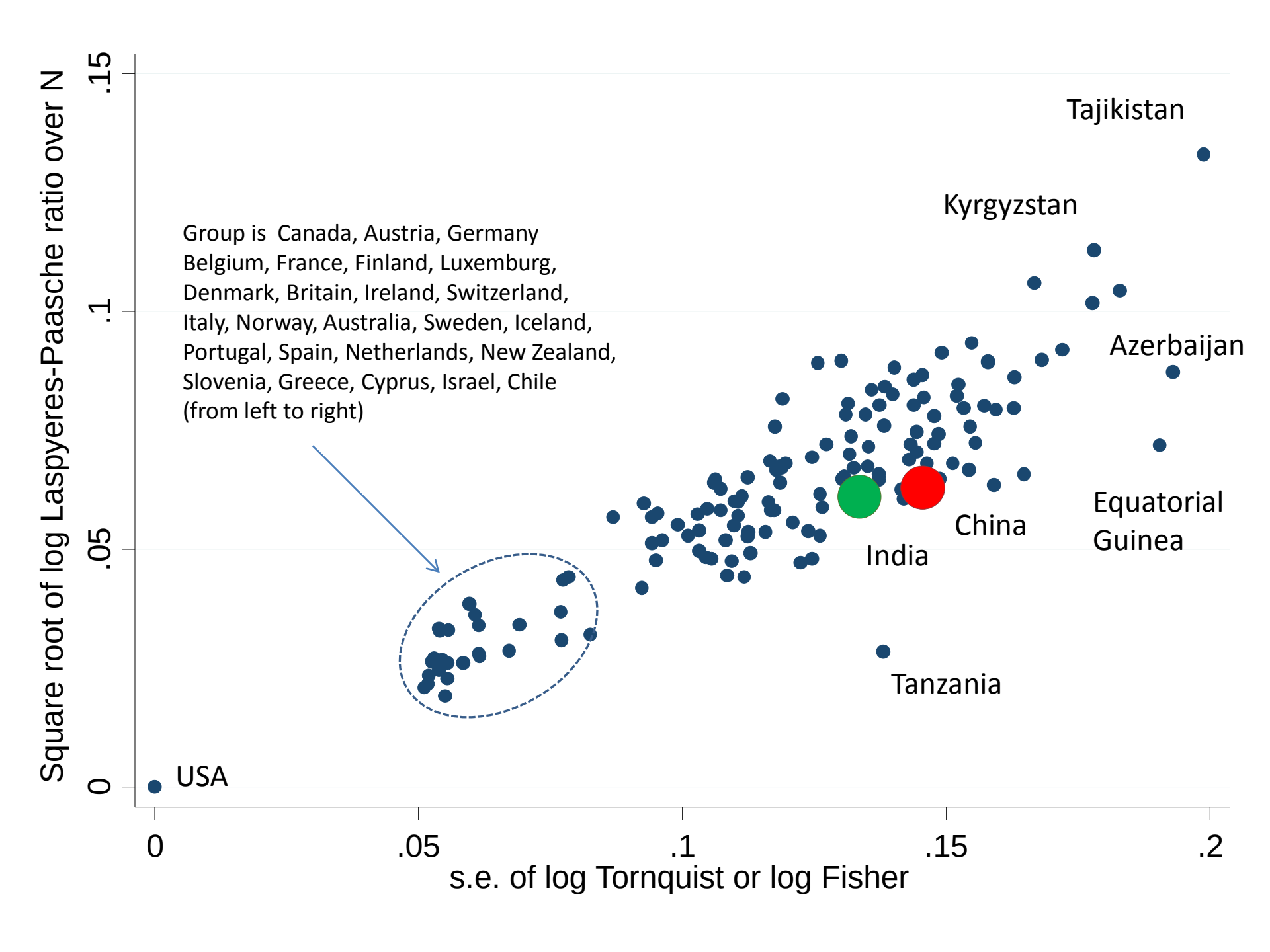
- Suppose B is China, and A is the US, that the share of non-traded goods is larger in the US, and the relative price of non-traded goods is rising more rapidly in China (BS), then RHS is *positive*, and PPP will rise over time relative to the updating formula

Does this do it?

- About a quarter of the divergence between CPIs and PPPs
- This bias can be avoided in future if the ICP updates at a finer level of disaggregation
- The rest remains a puzzle: possibly just random noise
  - good to have more people thinking about this:
- Other more general issues, about the precision of the ICP, PWT, and other numbers based on them
  - Large standard errors? Need a concept
- Richard Stone (1949)
  - "Why do we need to compare the U.S. with, say, India or China? Everybody knows that one country is very rich and another country very poor, does it matter whether the factor is thirty or fifty or what?"
- We have learned a lot since, or have we?

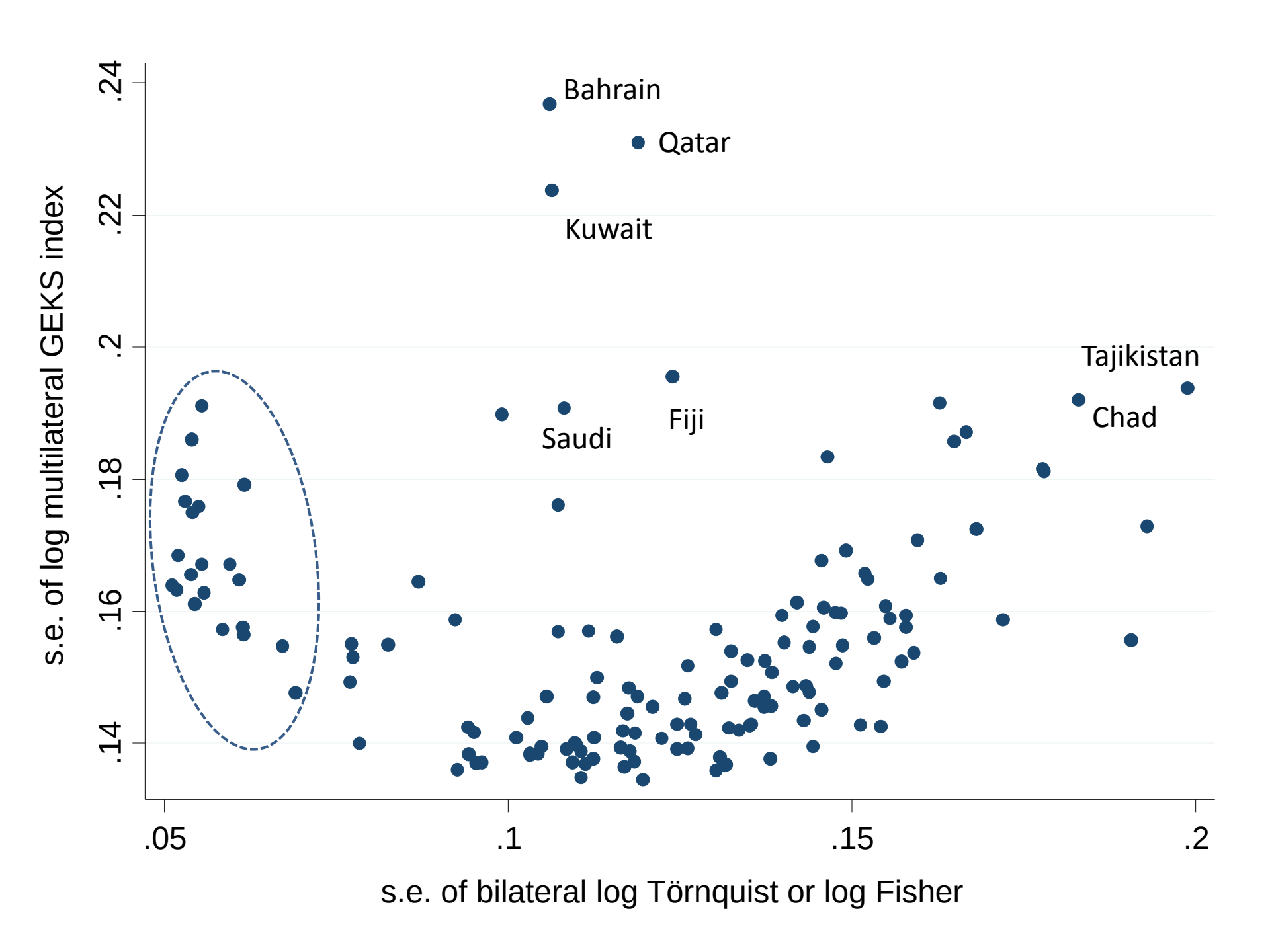
Precision of PPP rates

- Compared with a "star" country, e.g. US, each country has a N-vector of price relatives, one for each good
- Variance of log price ratios between pairs of countries source of PPP index uncertainty
- Theorem: log of Laspeyres to Paasche ratio for A relative to B (divided by N) is approximately equal to the variance over goods of log price ratios for A to B
- Standard error of PPP comes from thinking about log price ratios as drawn from a distribution
  - PPP distribution calculated using (fixed) weights applied to price draws


Square root of log Laspyeres-Paasche ratio over N
0
.05
.1
.15
0
.05
.1
.15
.2
s.e. of log Tornquist or log Fisher
Group is Canada, Austria, Germany
Belgium, France, Finland, Luxemburg,
Denmark, Britain, Ireland, Switzerland,
Italy, Norway, Australia, Sweden, Iceland,
Portugal, Spain, Netherlands, New Zealand,
Slovenia, Greece, Cyprus, Israel, Chile
(from left to right)
Tajikistan
Kyrgyzstan
Azerbaijan
Equatorial
Guinea
China
India
Tanzania
USA

Notes on the figure

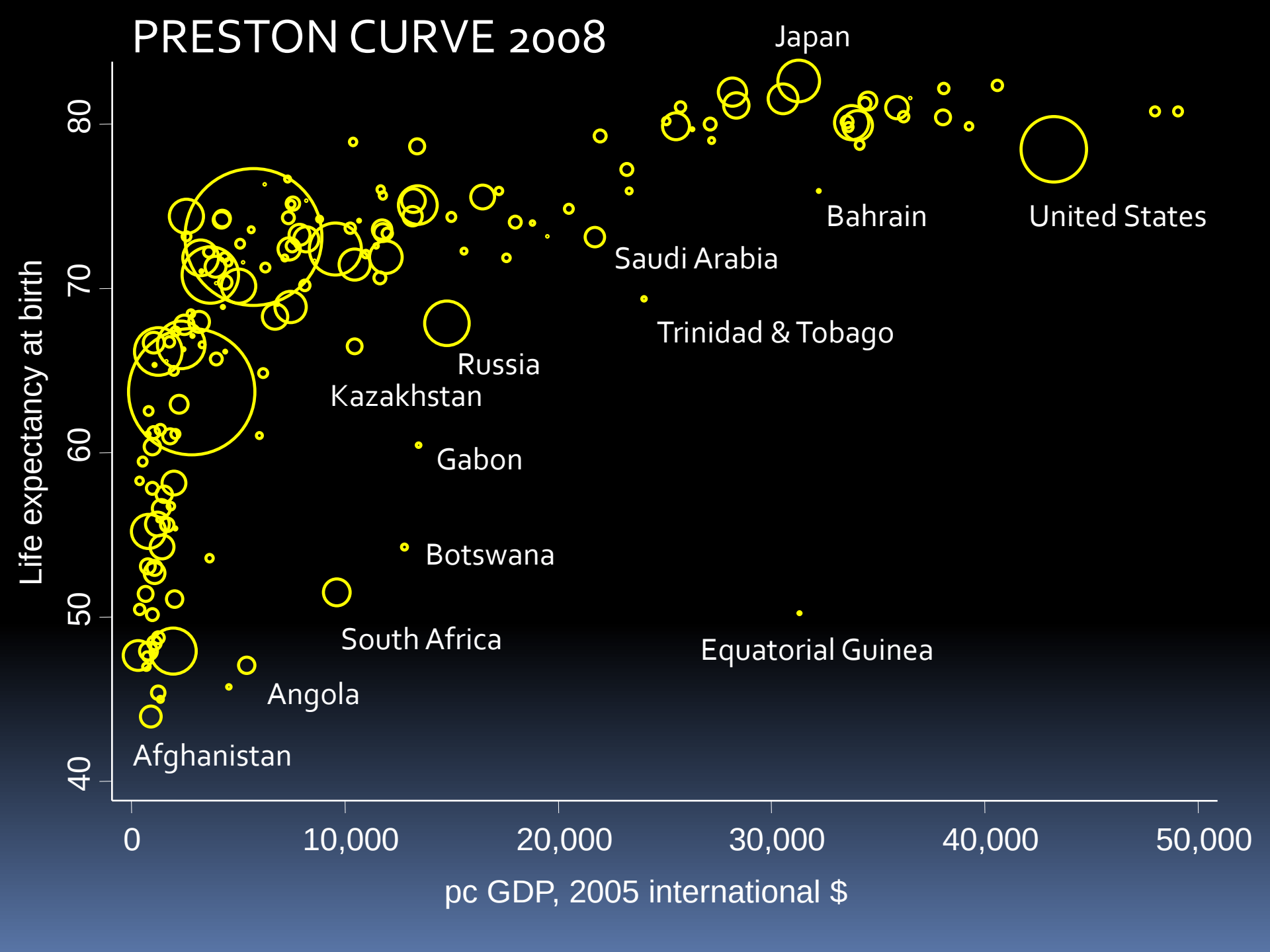
- The standard errors are *large*
  - 2 s.e. for China and India is around 30 percent
- Much smaller for the group on the left
  - But still substantial, ten percent
- Stone again in 1949
  - "I do not expect a very rapid resolution of the intellectual problems of making welfare comparisons between widely different communities"
- But ICP indexes are multilateral, not bilateral, and we need standard errors for those
  - Essence of multilateral is that transitivity is enforced, and bilateral indexes are not transitive
  - So what happens?


Bahrain
Qatar
Kuwait
Saudi
Fiji
Tajikistan
Chad
.24
.22
.2
.18
.16
.14
s.e. of log multilateral GEKS index
.05
.1
.15
.2
s.e. of bilateral log Törnquist or log Fisher

Multilateral s.e.'s

- Multilateral standard errors are typically larger
  - Average 15 percent instead of 12 percent
- Dispersion of ML standard errors smaller
  - Transitivity is spreading the errors
  - Poor bilateral is buttressed by ML comparisons
- Close countries have *much* larger s.e.'s
  - ML is a bad idea for them
  - Bringing Tajikistan into the Canada US comparison is not necessarily a good idea
- Middle group of countries where costs of transitivity are balanced by the gains
- Still substantial uncertainty, big standard errors

HEALTH AND INCOME


PRESTON CURVE 2008
Japan
Bahrain
United States
Saudi Arabia
Trinidad & Tobago
Russia
Kazakhstan
Gabon
Botswana
South Africa
Equatorial Guinea
Angola
Afghanistan
Life expectancy at birth
40
50
60
70
80
0
10,000
20,000
30,000
40,000
50,000
pc GDP, 2005 international $

Correlations

- Health and income are positively correlated
  - Across countries
  - Within countries
- Reasons are subject to continuing debate
- Looking at both together, ideally within as well as between countries
  - Shows more inequality in the world for combined measures
  - Changes patterns of growth, more for poorer countries
- Measurement tools are extensions (and improvements) of Sen's 1973 suggestion of
  - LE*income*(1-gini)
- Important and insightful work by Fleurbaey using money metric utility theory, by Becker et al, and especially comprehensively by Jones and Klenow using EU theory

LE and wellbeing?

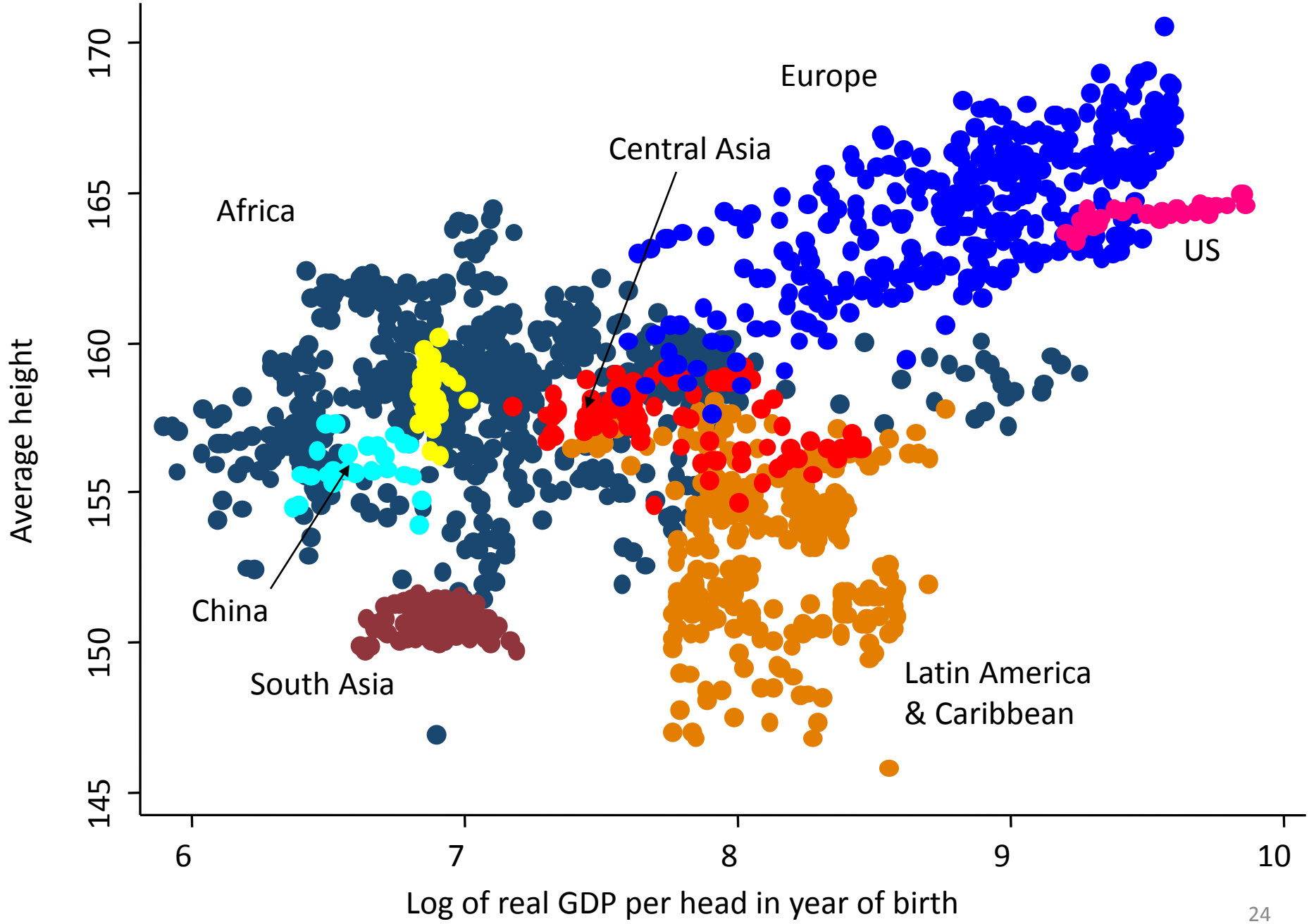
- Rapid worldwide decreases in mortality since WW2
- Different places in the age distribution on either side of the epidemiological transition
  - Poor countries, saving kids, from ARIs, diarrhea, and vaccine preventable disease
    - Some cases of LE increasing at 3 years per year
  - Rich countries saving middle-aged or elderly adults from smoking and from heart disease (and cancer)
  - Convergence in rich countries is so strong, got to be smoking and health care innovations (antihypertensives)
- What are the welfare implications of these different age structures of mortality decline?
  - Can allow for different ages, as JK do
  - But there are further issues if fertility adapts

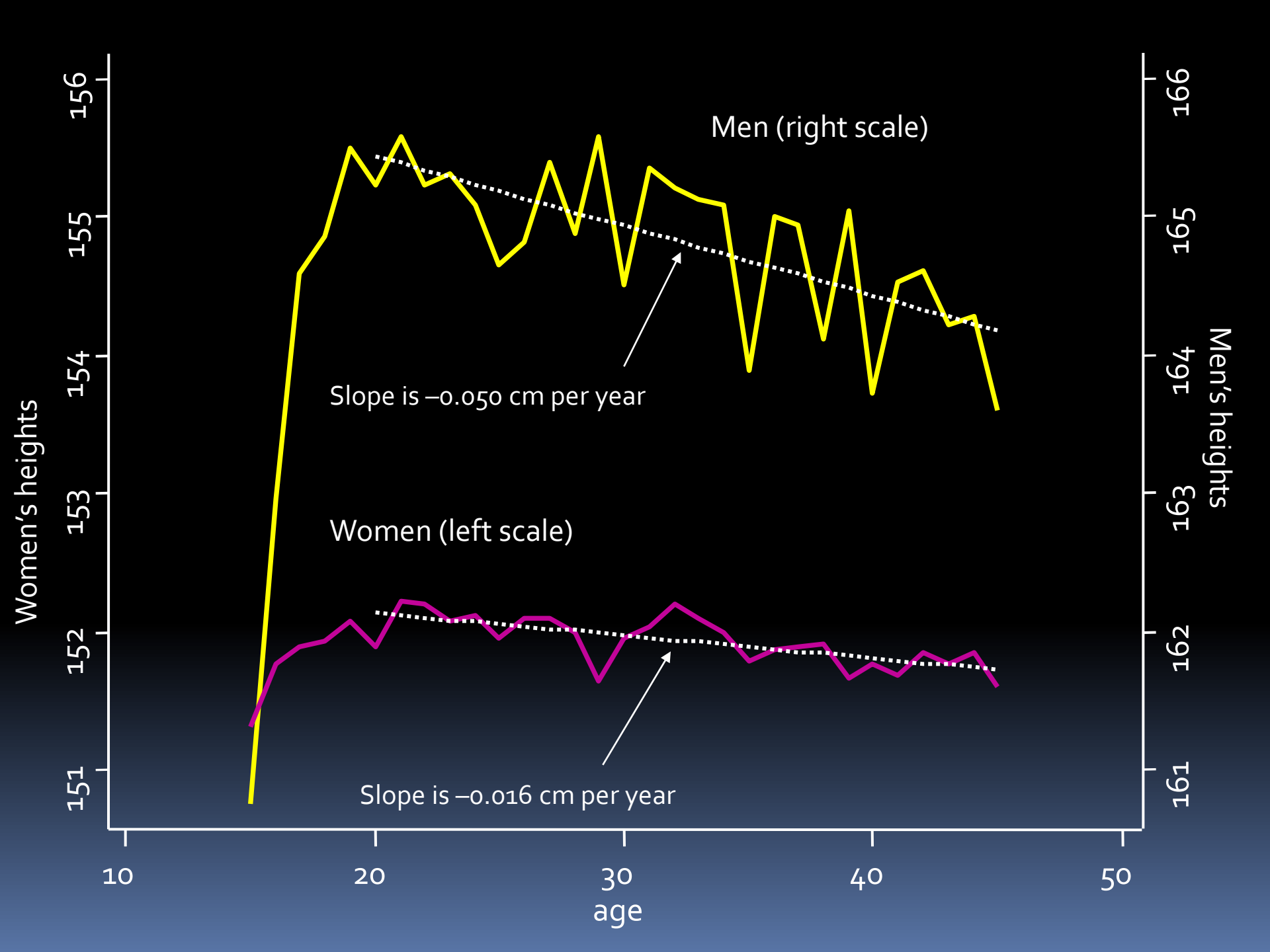
Mortality and fertility

- (A) Suppose each woman bears 6 kids: 3 die at birth, 3 survive until 60
  - LE at birth is 30 years
- (B) UNICEF comes along and makes things better, so only 2 die, and 4 survive
  - LE is now 40 years
- (C) Families adapt to have only 4.5 kids: 1.5 die at birth, 3 survive until 60
  - LE is still 40
- Kids who would previously ceased to exist just after birth in A now cease to exist just before birth in C: welfare gain?
- Age structure of population in (C) is identical to age structure in (A), survival rates of everyone alive are the same
- Gain in WB is the reduction in fertility burden for women (not watching infants die), not the gain in LE

Heights

- Adult height is largely determined in childhood by combination of material resources and health
  - Combination is not necessarily of welfare significance
- Adult height does not change until around age 50
  - Tree ring property: a single cross section can be used to read history of the health/income combination in childhood
    - Komlos "biological standard of living"
- Not a perfect indicator of morbidity, even in childhood, but we have few "non-fatal" measures of population health
  - Taller people are on average smarter and earn more
  - Cognitive function: brain and body develop together
- Changing patterns of national heights


Average height
170
165
160
155
150
145
6
7
8
9
10
Log of real GDP per head in year of birth
Europe
Central Asia
Africa
US
China
South Asia
Latin America
& Caribbean


Men (right scale)
Slope is –0.050 cm per year
Women (left scale)
Slope is –0.016 cm per year
Women's heights
Men's heights
age
151
152
153
154
155
156
161
162
163
164
165
166
10
20
30
40
50

Facts about heights

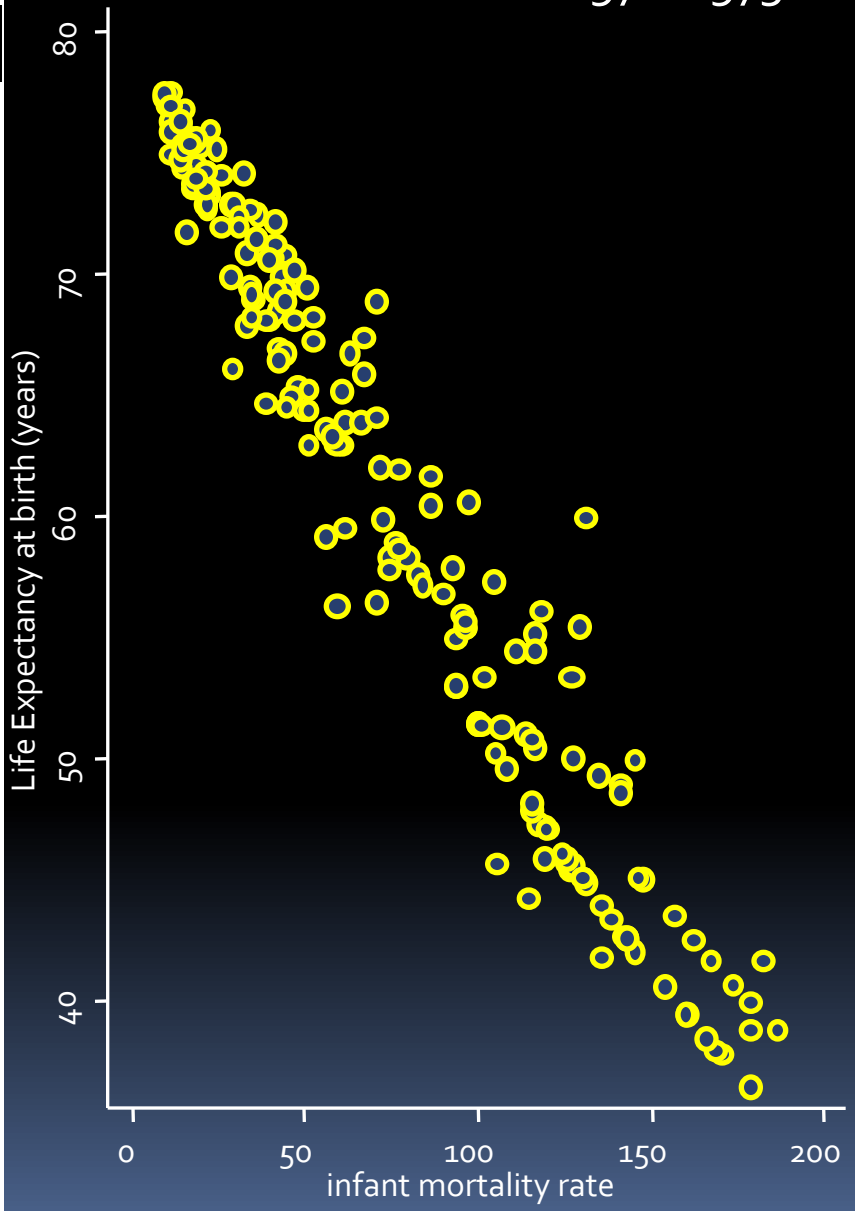
- Apart from now rich countries, no correlation between pc GDP in year of birth and adult height
  - In rich countries for about 200 years, heights have increased at about 1 cm per decade
  - In China now about the same
  - In India now, about half that, and less still for women
    - Slow growth of men may be attributable to discrimination against women
  - In Africa, women have been shrinking (Subramanian et al) at least during the years of negative growth
- Correlation between height and GDP per head is increasing
  - Random initial distribution from ecological niches
  - Increases related to GDP growth and nutrition as markets integrate
  - Relation between GDP and height is becoming closer over time

INTERNATIONAL DATA

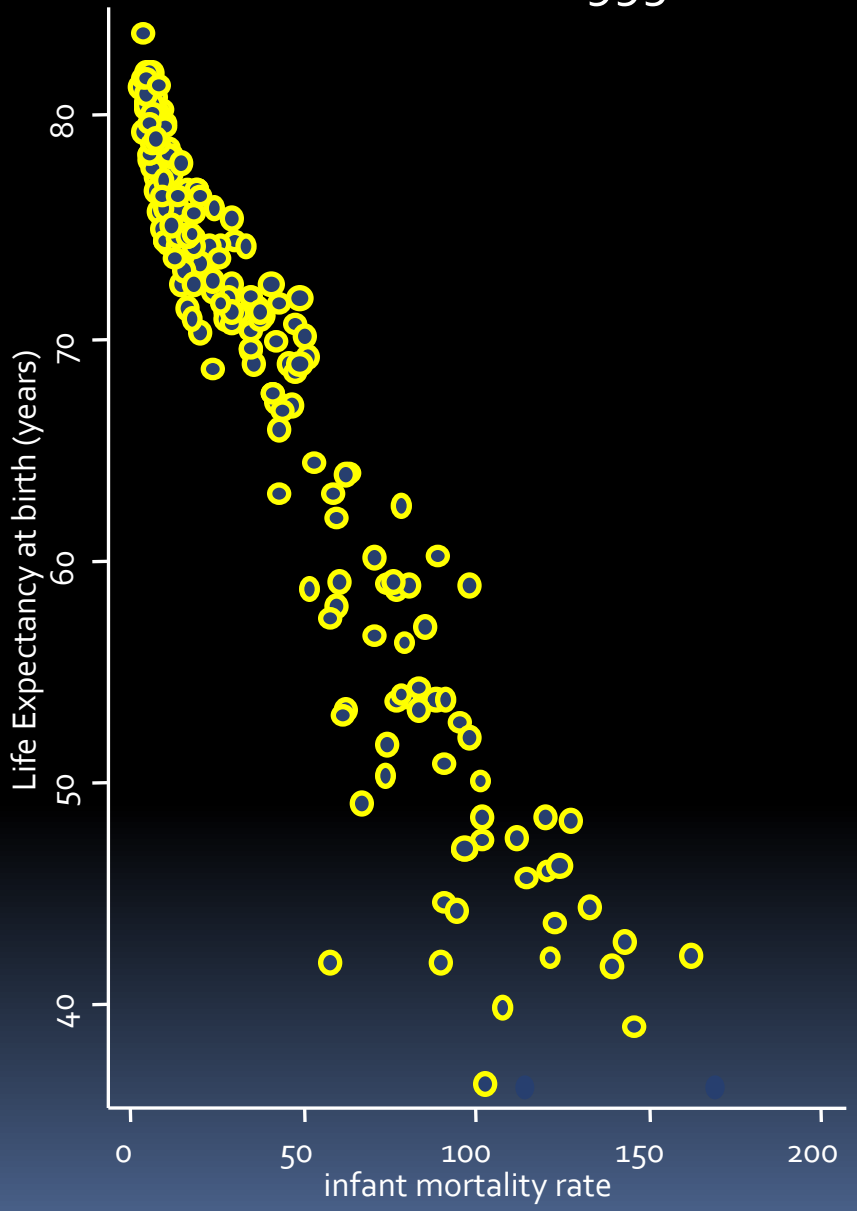
Bricks without straw

- UN system accepts national estimates of GDP
  - PPP only deflates them by price indexes
- Household surveys are inconsistent with national accounts
  - Different growth rates of mean income/consumption
  - Allows much mischief with poverty estimates
- Disasters in FAO hunger measures
- World Bank arbitrarily changes global poverty line, and demands renewed attention to poverty
- Large well-funded industry of making up numbers
  - Mortality rates, by cause of death, for countries with no vital registration
  - Including doctors "reconciling" different PPP measures of GDP
  - GDP for countries that didn't exist
- Mortality rates are not measured in much of the world

LE v IMR UN data 1970-1975



LE v IMR UN data 1995-2000



Politics of data

- Foundations or UN system will not pay for data collection
  - They believe in *doing* things, even doing them blind
  - No incentives for them to have good data
- No international control, in contrast to often elaborate domestic controls
  - Compare US CPI and unemployment
  - Perhaps high quality global statistics, like global justice, is a cosmopolitan fantasy
  - Certainly no political constituency that will punish egregious errors/use of data for self-promotion
    - Who needs these numbers?

HAPPINESS & DEVELOPMENT

Happiness measures

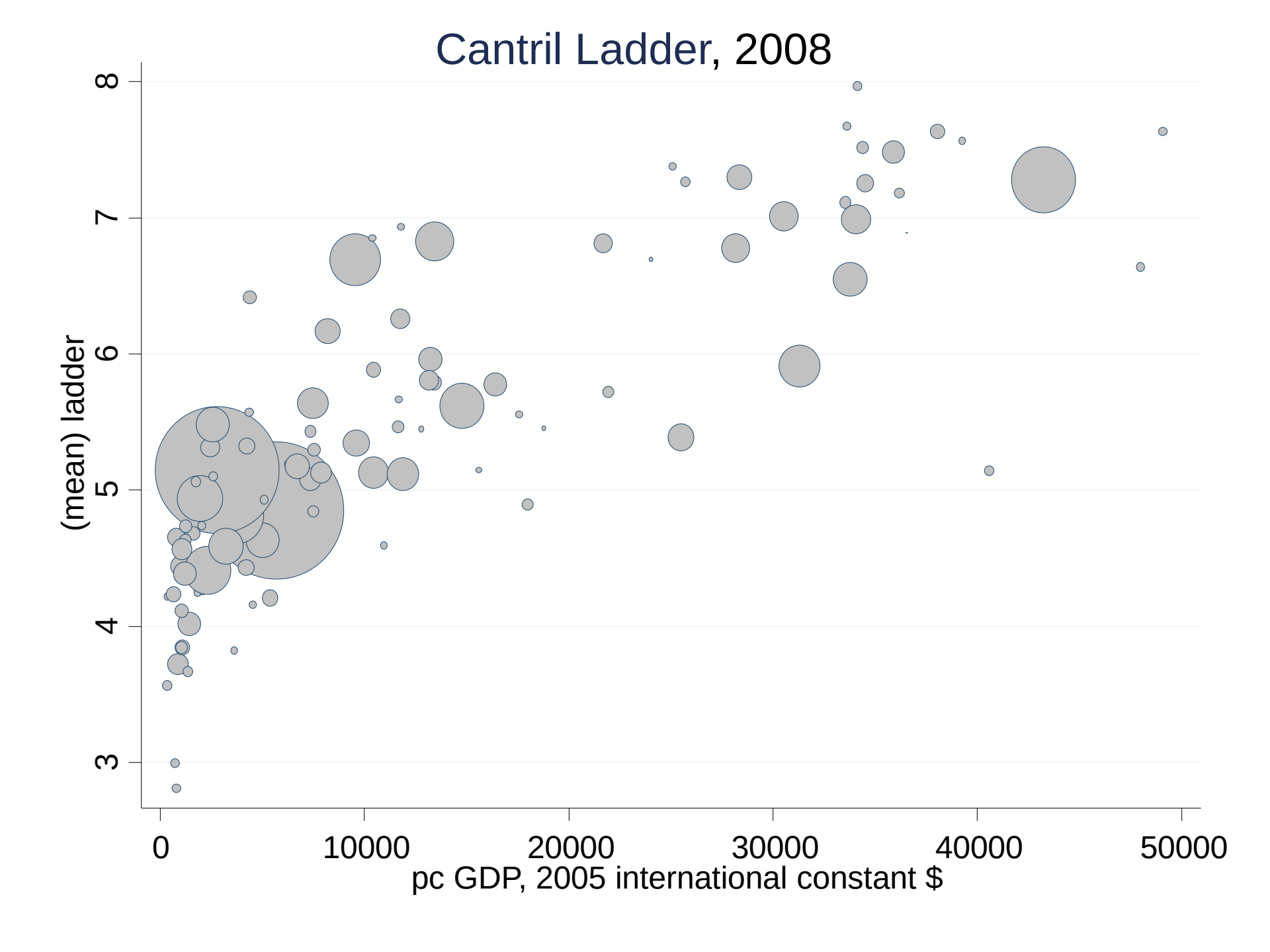
- Much enthusiasm, but many concerns
  - Discomfort with making interpersonal comparisons
  - Do self-reported questions mean anything?
- Adaptation problem
  - If people get used to being deprived and oppressed, does that make it OK? (Sen, Nussbaum)
  - Happiness measures may ignore such deprivation
- Empirical evidence relevant but not dispositive
  - What are the correlates of SWB measures?
  - Is it tenable to base policy *only* on self-reported happiness?
    - Layard's version of Bentham
  - Or is life happy because it is full and rich, even though it may often involve pain and loss (Nussbaum)

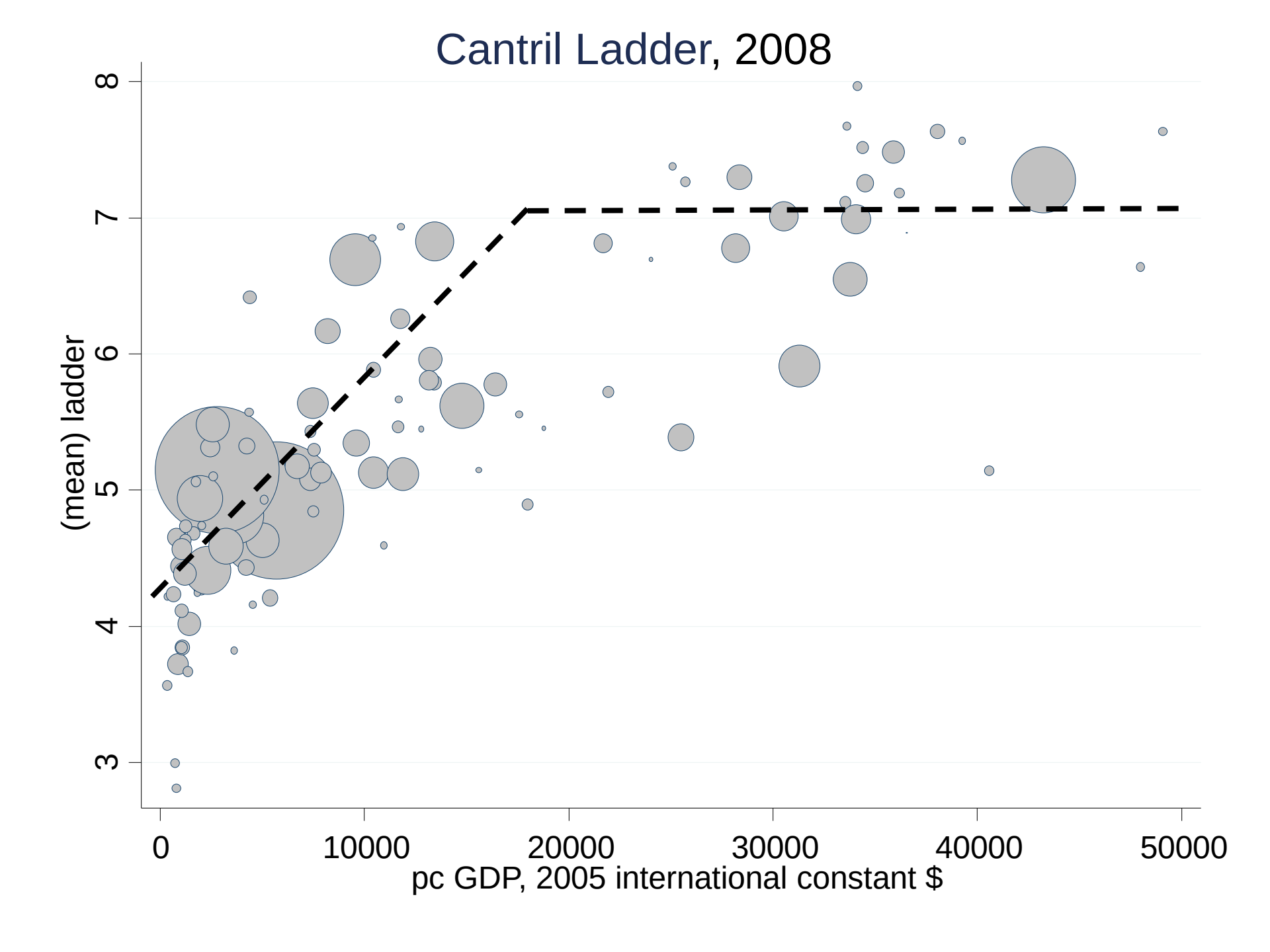
New information

- Gallup's World Poll samples all of the citizens of the world
  - 1,000 people in each country in each year
  - 160 countries now covered, since 2006
  - Identical questionnaires
- Data allow examination of many questions about SWB
  - Kahneman has been advisor since the beginning of these projects
- Are richer countries better off? Do people adapt to their conditions? Does growth improve the human lot ? (Easterlin)

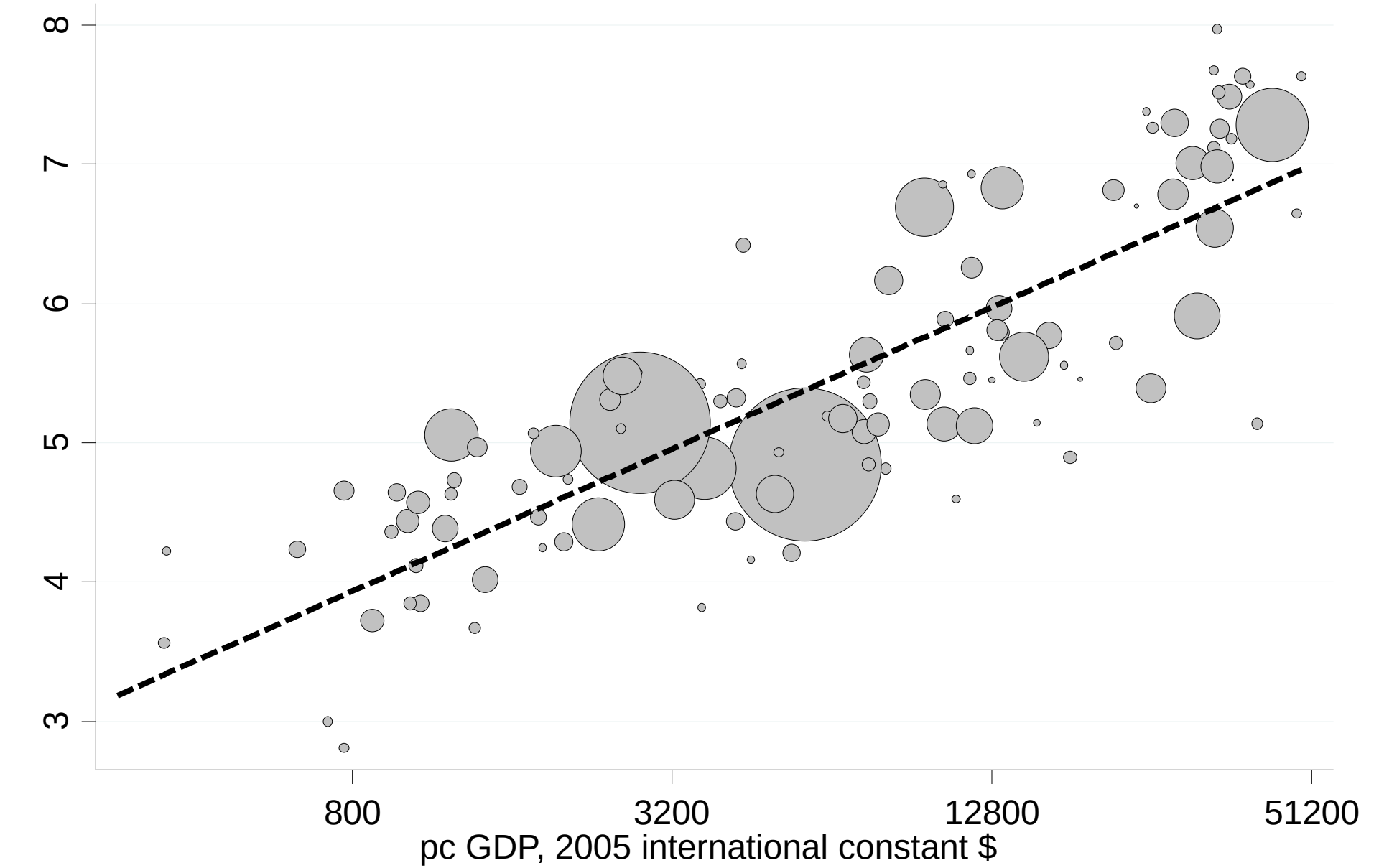
Cantril's ladder question

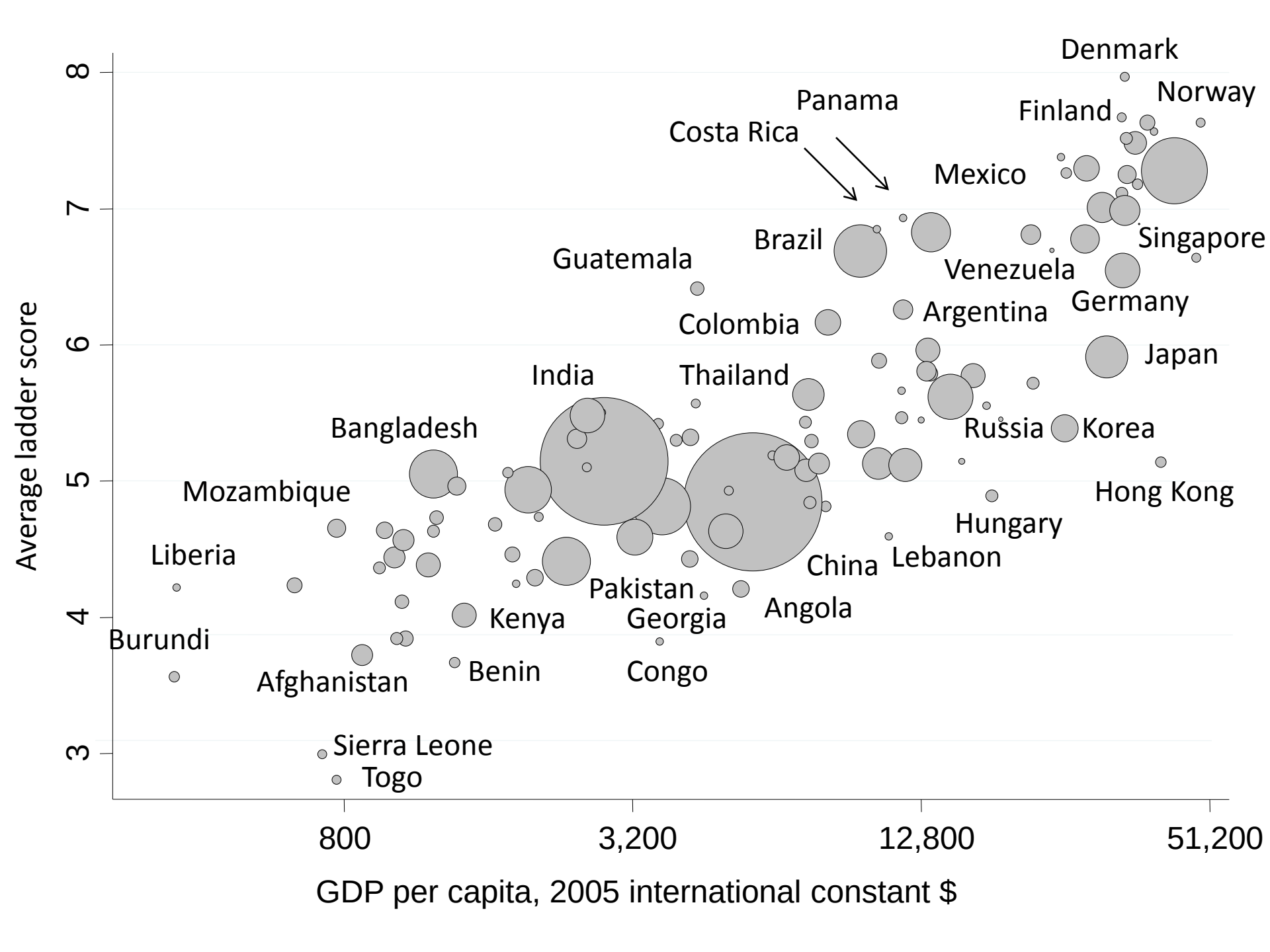
- *Please imagine a ladder, with steps numbered from 0 at the bottom to 10 at the top. The top of the ladder represents the best possible life for you and the bottom of the ladder represents the worst possible life for you. On which step of the ladder would you say you personally stand at this time?*
- Life evaluation measure, like life satisfaction, but hedonically neutral
- *Not* happiness
  - Happiness proper (mood, hedonic experience)
  - Did you experience a lot of happiness yesterday
- Distinction between *experiencing* life (hedonics, like happiness, sadness, enjoyment, etc) and *thinking about* life is crucial, conceptually and empirically
  - Different correlates, different adaptation, etc.


Cantril Ladder, 2008
(mean) ladder
3
4
5
6
7
8
0
10000
20000
30000
40000
50000
pc GDP, 2005 international constant $


Cantril Ladder, 2008
(mean) ladder
3
4
5
6
7
8
0
10000
20000
30000
40000
50000
pc GDP, 2005 international constant $

Cantril Ladder, 2008



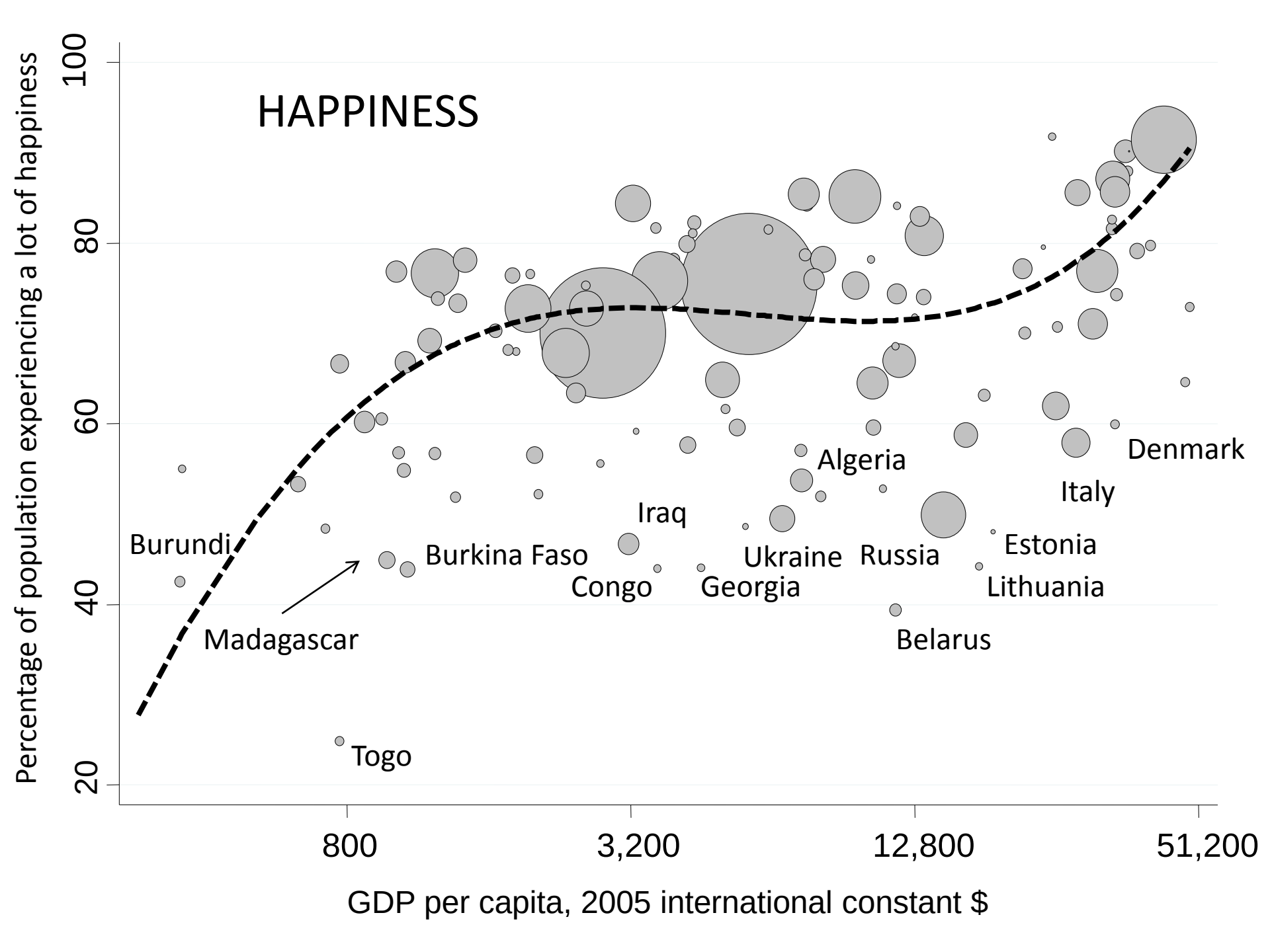

Denmark
Norway
Finland
Panama
Costa Rica
Mexico
Singapore
Brazil
Guatemala
Venezuela
Germany
Argentina
Colombia
Japan
India
Thailand
Bangladesh
Russia
Korea
Mozambique
Hong Kong
Hungary
Liberia
China
Lebanon
Pakistan
Angola
Kenya
Georgia
Burundi
Benin
Congo
Afghanistan
Sierra Leone
Togo
8
7
6
5
4
3
Average ladder score
800
3,200
12,800
51,200
GDP per capita, 2005 international constant $

Continuing debates

- Some continue to hold to the "no relation among rich countries position"
  - Anti-materialism is very strong, especially outside of economics
- No one thinks money doesn't matter in poor places
- Big argument now is what happens over time within countries as they get richer
  - Original Easterlin paradox for US and Japan
  - Recent challenges by Stevenson and Wolfers
  - Easterlin does not agree: China particular bone of contention
  - Note that it takes hundreds of years of growth to get from Togo to the United States
  - Growth effects are *inherently* harder to find amidst other thing
  - My guess is that Stevenson and Wolfers are right
  - Easterlin too ready to combine inconsistent surveys

SWB & relative income

- School of thought, particularly in Europe, and non-economists, that it is only *relative* income that matters
  - Relative to reference group, co-workers or neighborhood
- The ladder data show that this is false across countries
  - Togolese and Danes cannot be (only) comparing themselves with their neighbors
- I believe this is true within countries too
  - Between area slopes identical to within area slopes
  - A misunderstanding of the data
  - Can discuss in Q & A


HAPPINESS
Percentage of population experiencing a lot of happiness
100
80
60
40
20
800
3,200
12,800
51,200
GDP per capita, 2005 international constant $
Burundi
Madagascar
Burkina Faso
Togo
Iraq
Congo
Georgia
Algeria
Ukraine
Russia
Belarus
Estonia
Lithuania
Italy
Denmark

Evaluating SWB

- Life evaluation much more useful than happiness or other hedonics
  - Latter components of wellbeing, but certainly not interpretable as utility
  - Development is not worth much by such measures
- Life evaluation is correlated with factors that we think ought to belong in WB but are not captured by income-based measures
- BUT: evidence that people do not know how to answer the question
  - Ask about the direction the US is going
  - Then ask about your own life
  - Latter contaminated by former
  - Only if people are reminded: otherwise doesn't matter what you think of the direction the US is going

A wellbeing dilemma

- Seem to be able to measure emotional states
  - Were you happy a lot yesterday?
  - But not what we want
- Life evaluation requires cognitive effort and is difficult
  - "reports of subjective well-being (SWB) do not reflect a stable inner state of well-being. Rather they are judgments that individuals form on the spot, based on information that is chronically or temporarily accessible at that point in time, resulting in pronounced context effects" Schwarz+Strack (1999)
- Evaluative measures have real content but unreliable
  - More interesting for things economics cares about
  - Can we fix them technically?
  - Are there circumstances where they are OK?
  - Lots remains to be done

THANK YOU!