FLS 6415 - Causal Inference for the Political Economy of Development

Week 1 - Introduction

Jonathan Phillips

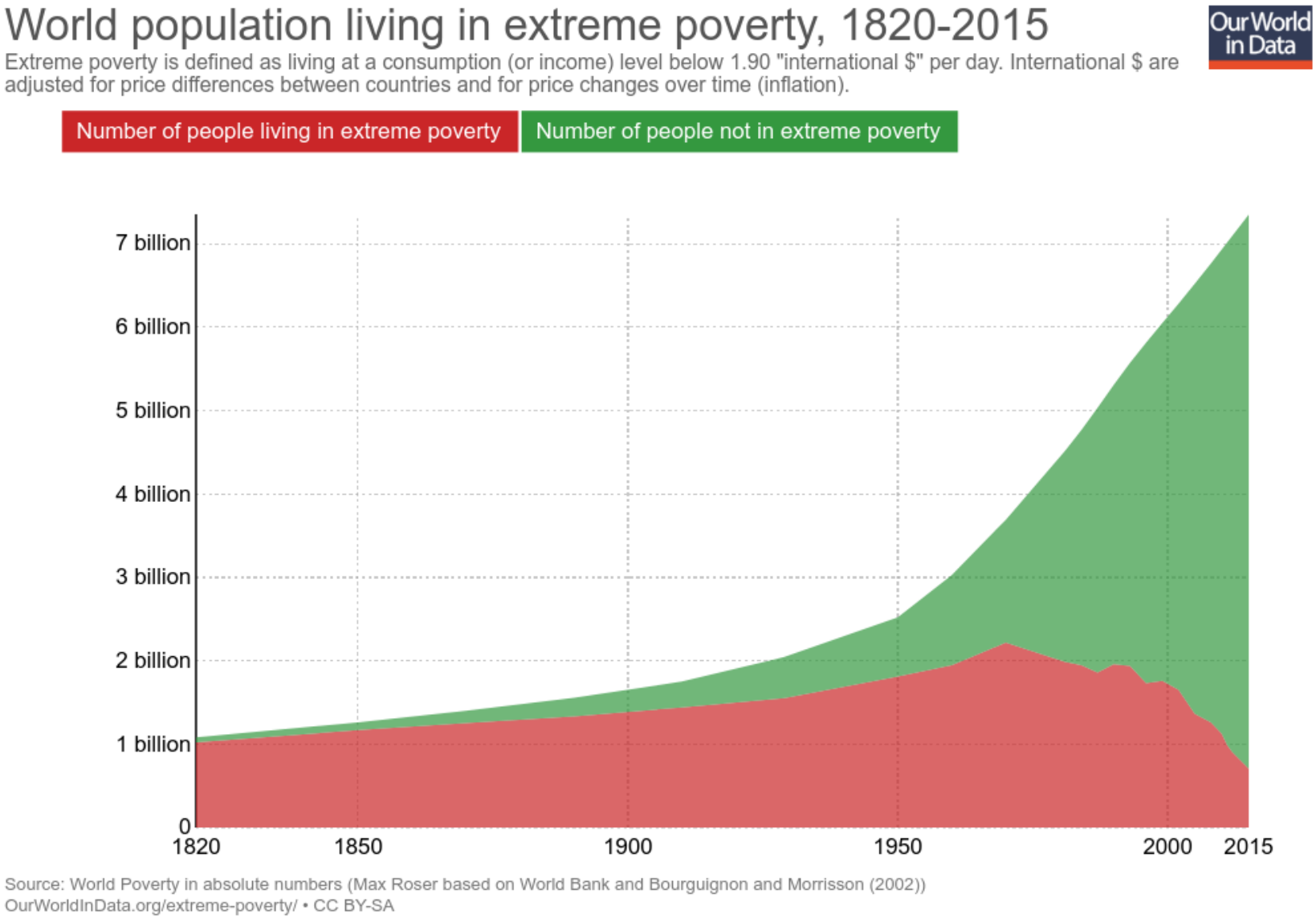
August 2017

The Big Picture

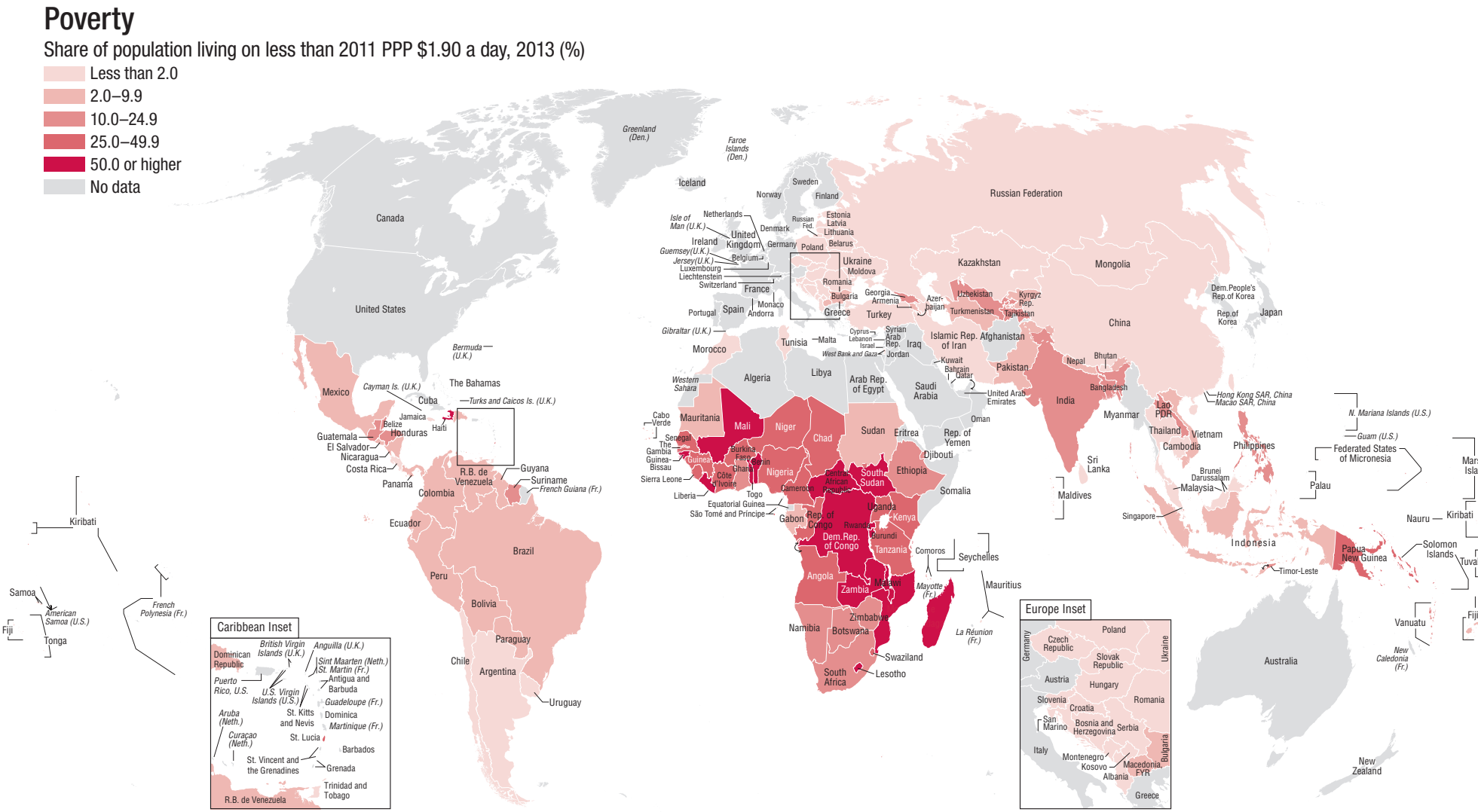
- **Why** are some places poor and some rich?
  - Many reasons! Geography, History...
  - We focus on **Politics**: Societies' decisions over who gets what
- **How** do we explain why some places are poor and some rich?
  - We focus on **causal inference**: How politics **causes** development outcomes

- By Development we mean Core Freedoms and Capabilities (Sen 1999)
- The freedom/capability to:
  - Avoid violence
  - Access economic opportunities (jobs)
  - Afford basic needs (shelter, food)
  - Be literate and numerate
  - Exercise political rights (free voting, free speech)

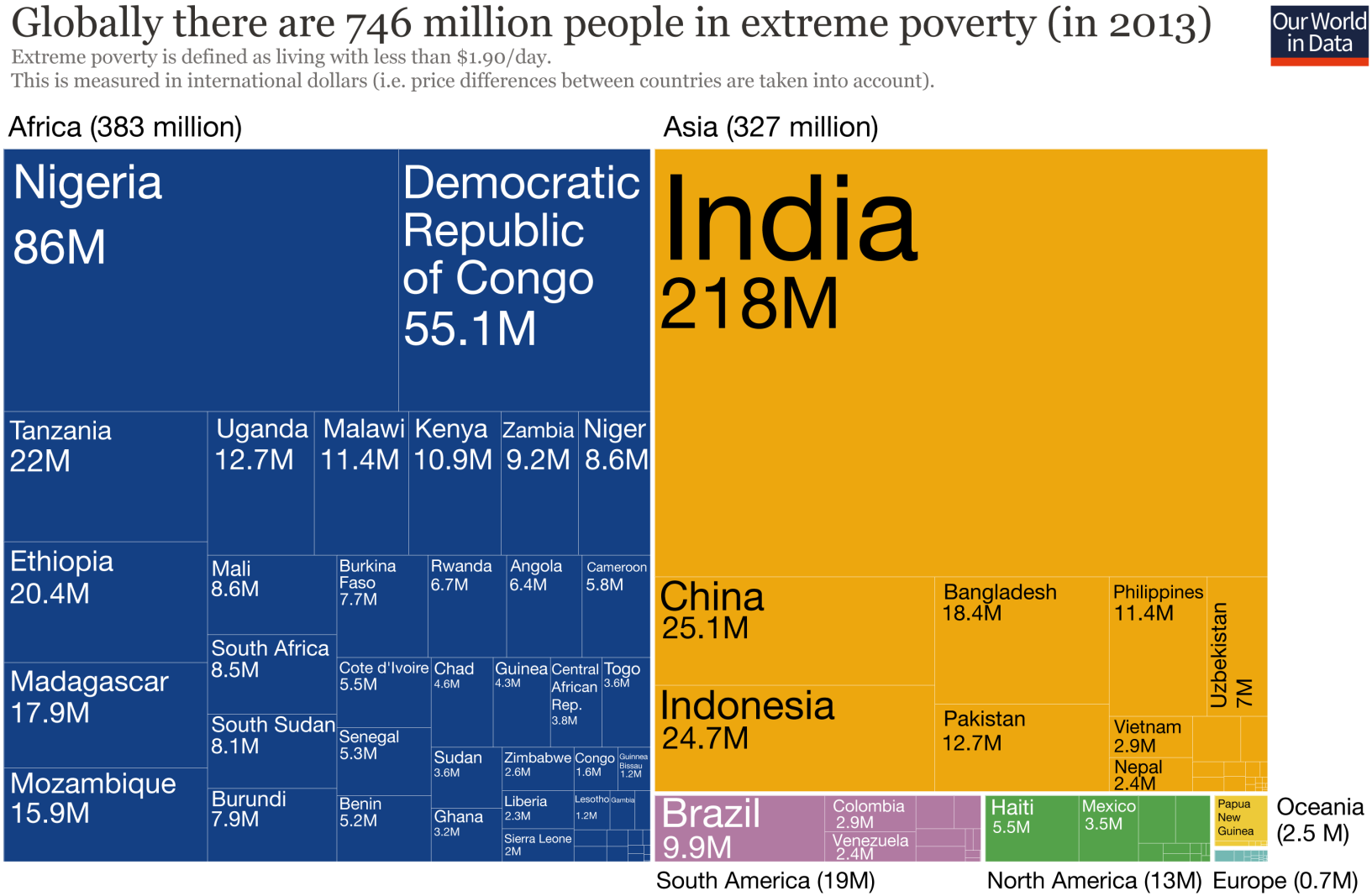
The Stylized Facts of Development



The Stylized Facts of Development

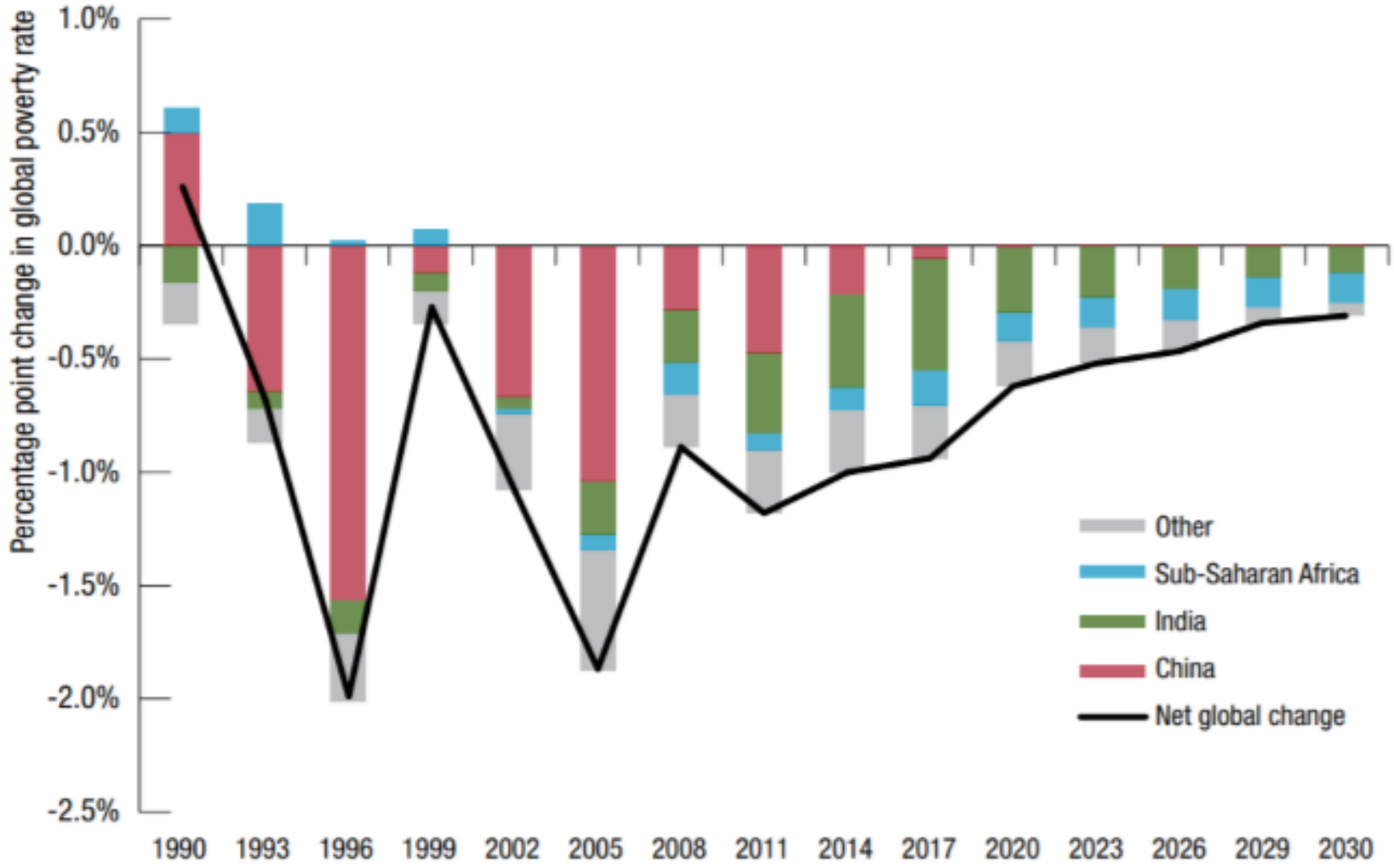


The Stylized Facts of Development

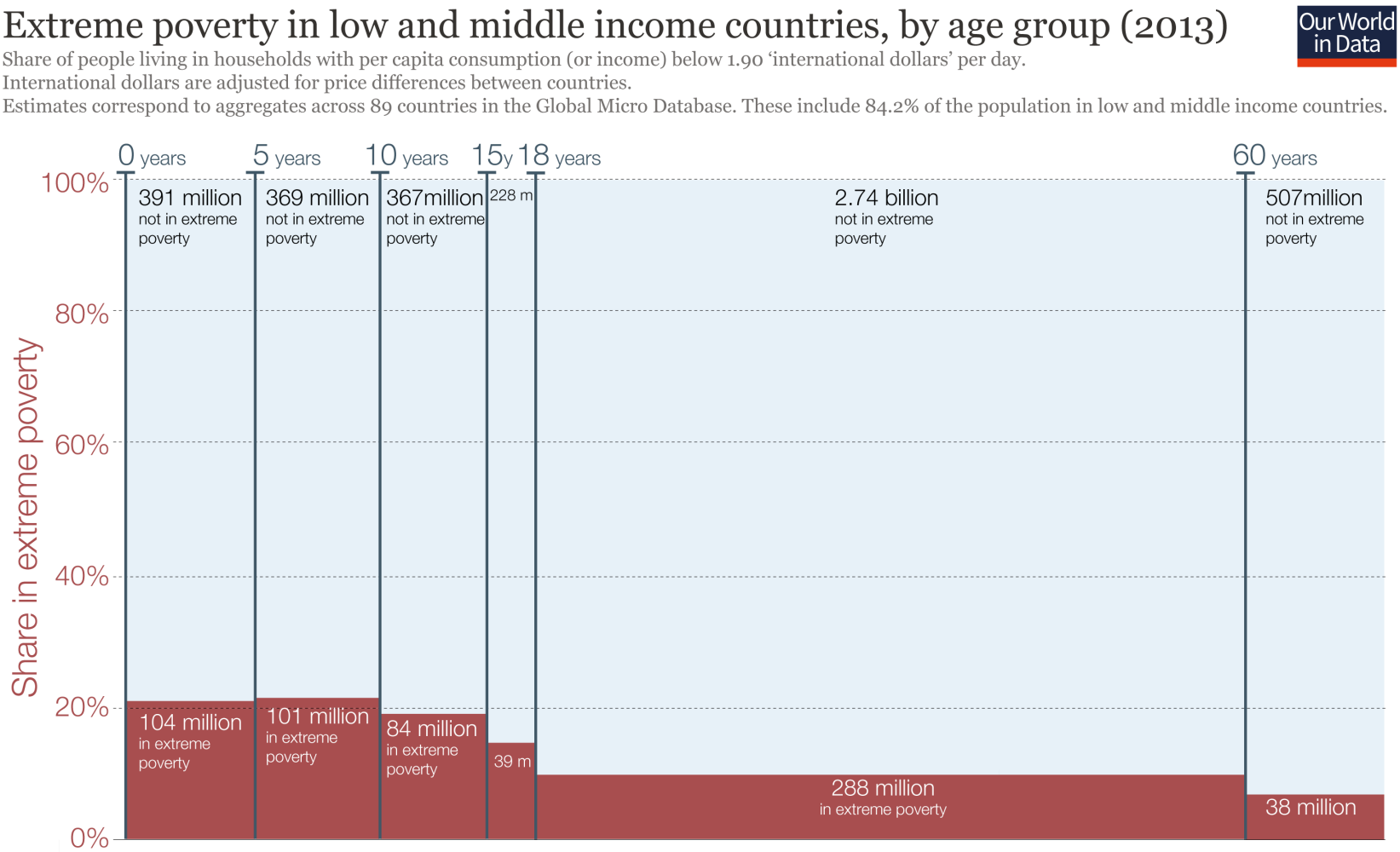


The Stylized Facts of Development

FIGURE 7: AVERAGE ANNUAL PERCENTAGE POINT REDUCTION IN GLOBAL POVERTY BETWEEN 1990 AND 2030 (OFFICIAL ESTIMATES AND BASELINE SCENARIO)

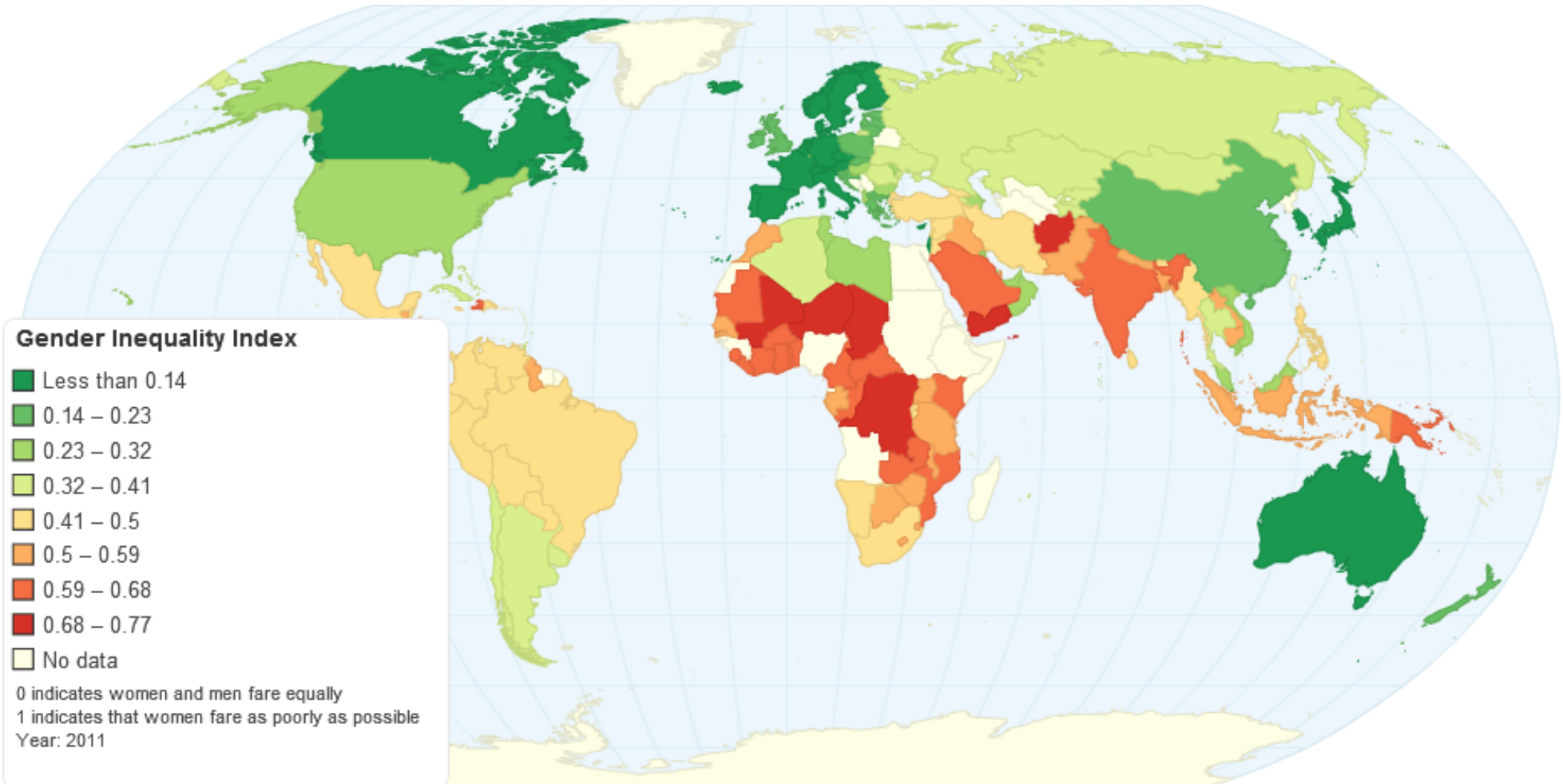


The Stylized Facts of Development

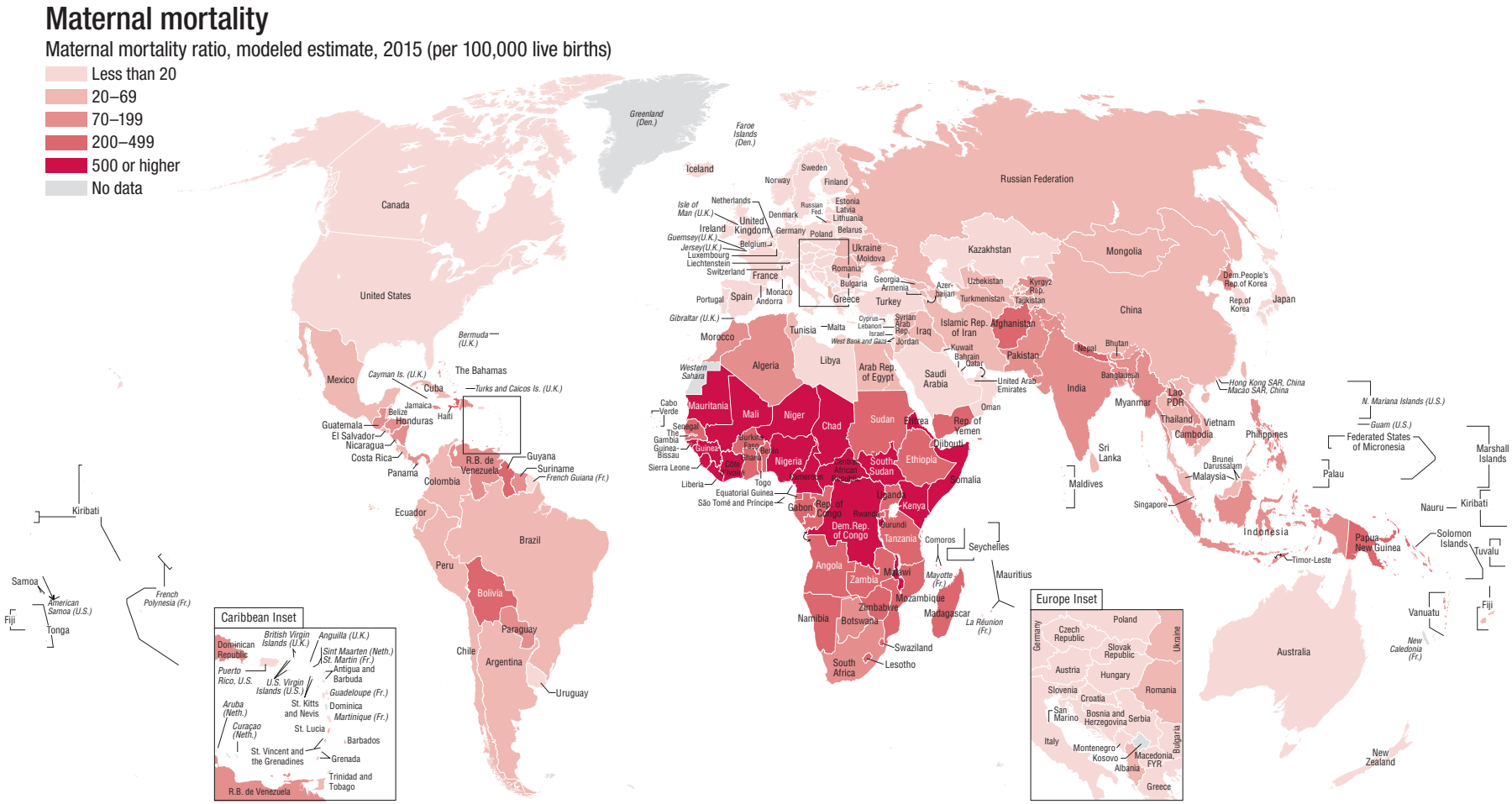


Data source: Newhouse, Suarez-Becerra, Evans, and Data for Goals Group (2016) – "New Estimates of Extreme Poverty for Children." Policy Research Working Paper 7845, World Bank
Data Note: Data comes from surveys taken between 2009 and 2014, but all figures are extrapolated to represent the estimates of extreme poverty in 2013.
The source defines the universe of low and middle income countries as all countries except: Australia, Belgium, Cyprus, Finland, France, Germany, Greece, Iceland, Ireland, Israel, Italy, Luxembourg, Japan, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom and United States.
This data visualization is available at OurWorldinData.org. There you find more visualizations and research on extreme poverty.

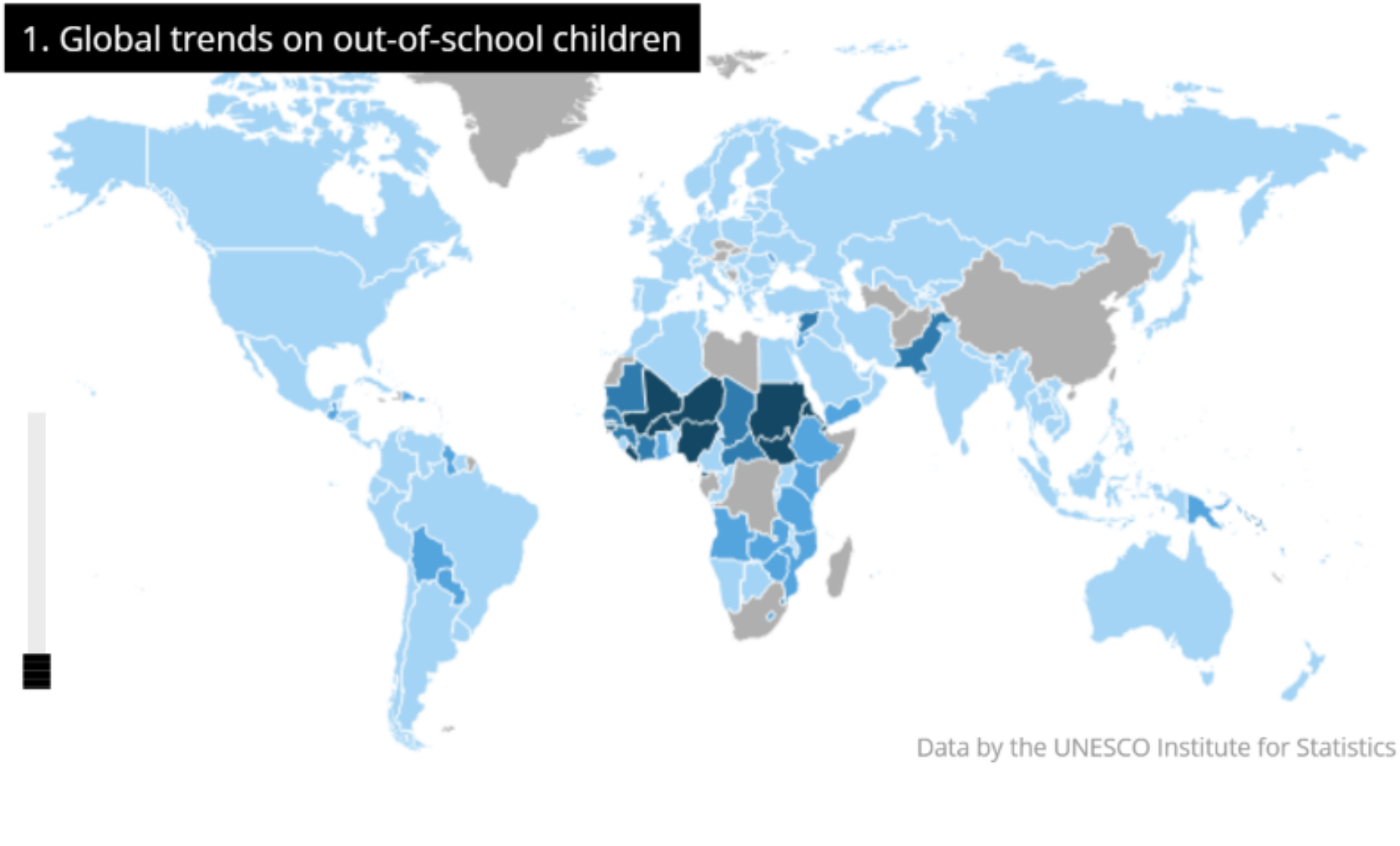
The Stylized Facts of Development



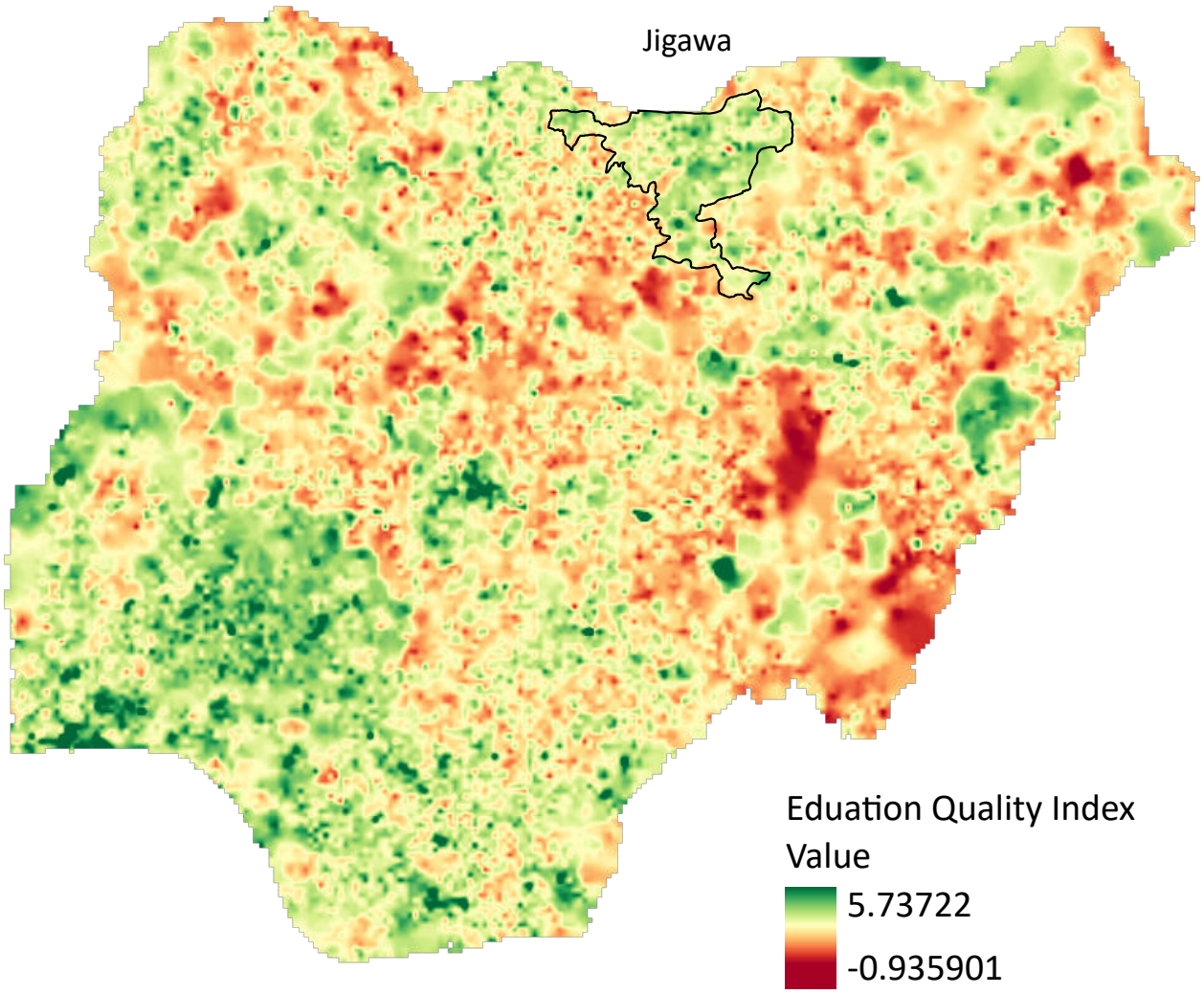
The Stylized Facts of Development



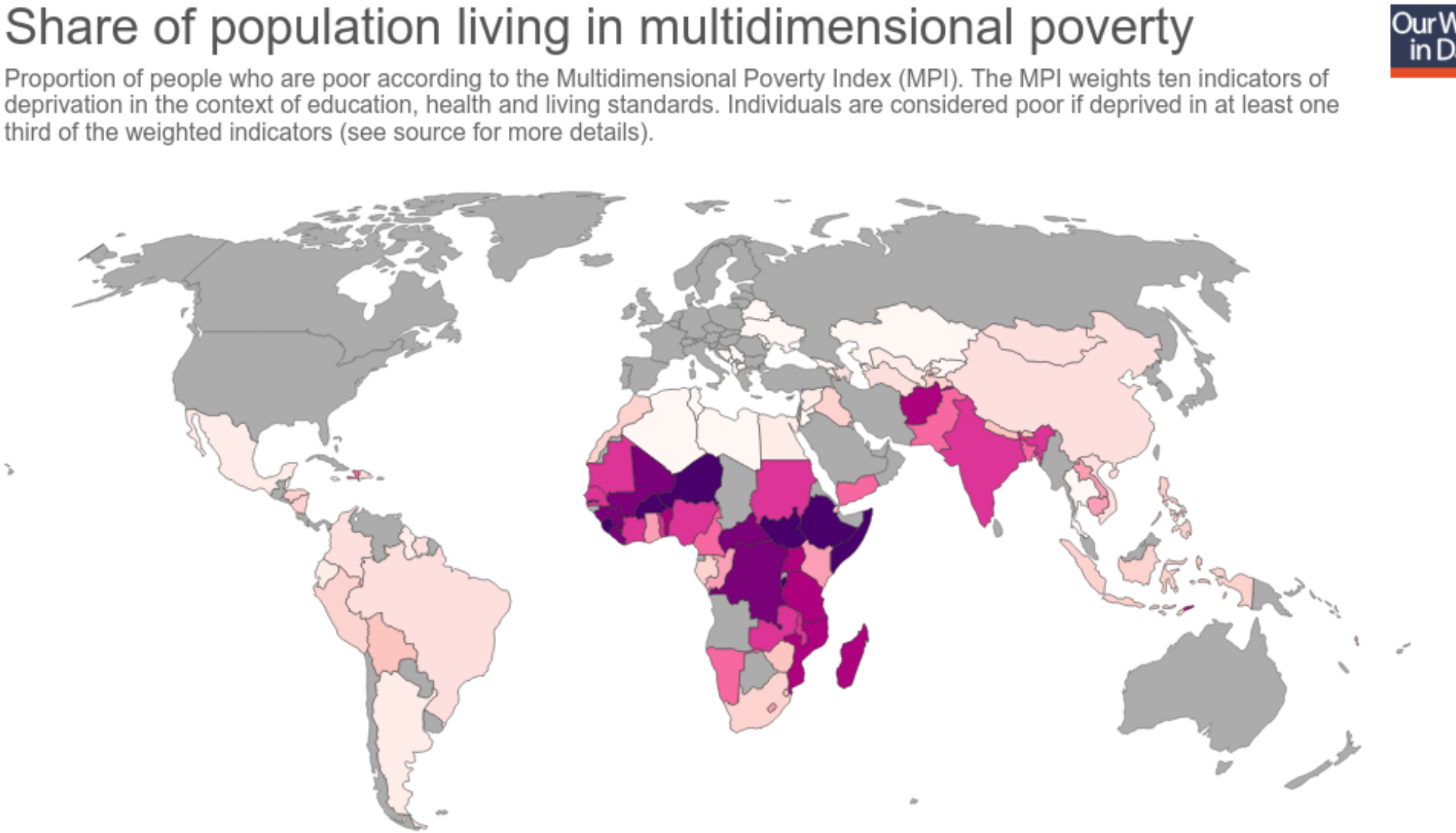
The Stylized Facts of Development



The Stylized Facts of Development

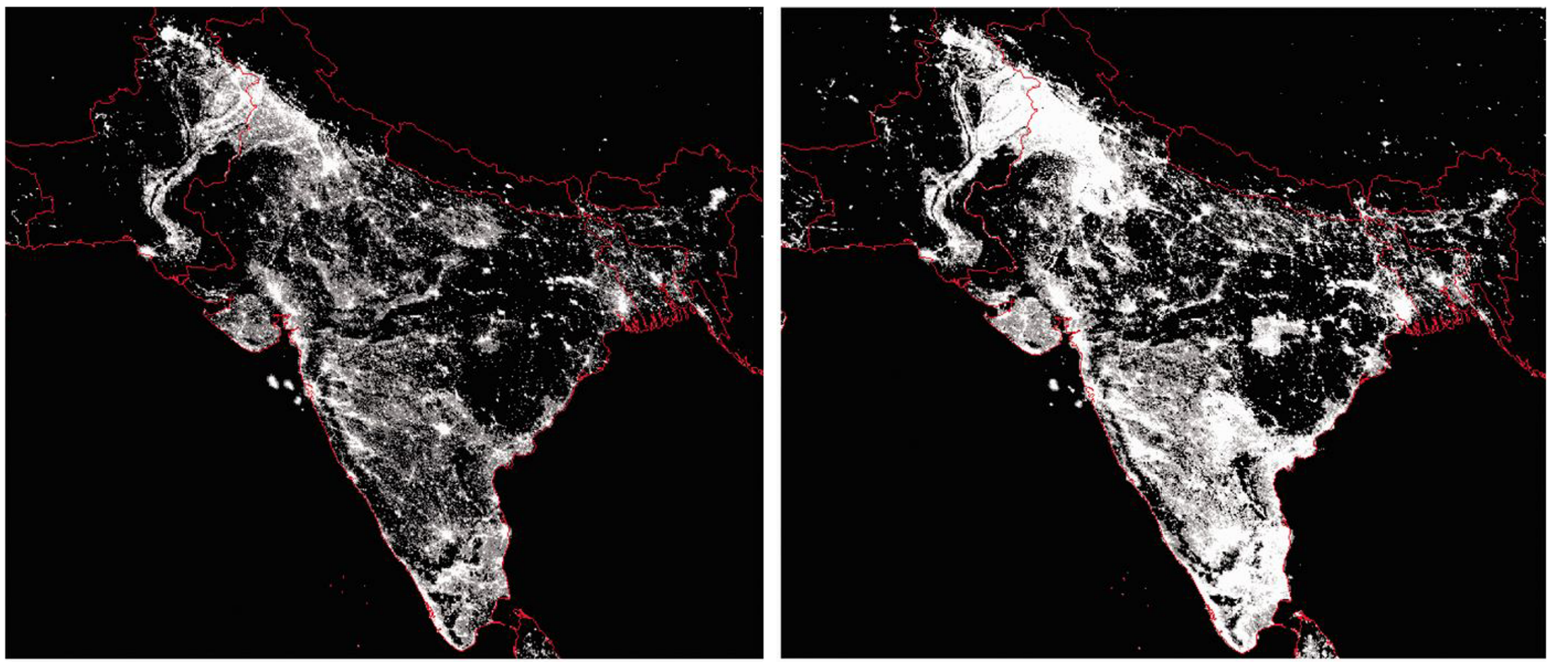


The Stylized Facts of Development



The Stylized Facts of Development

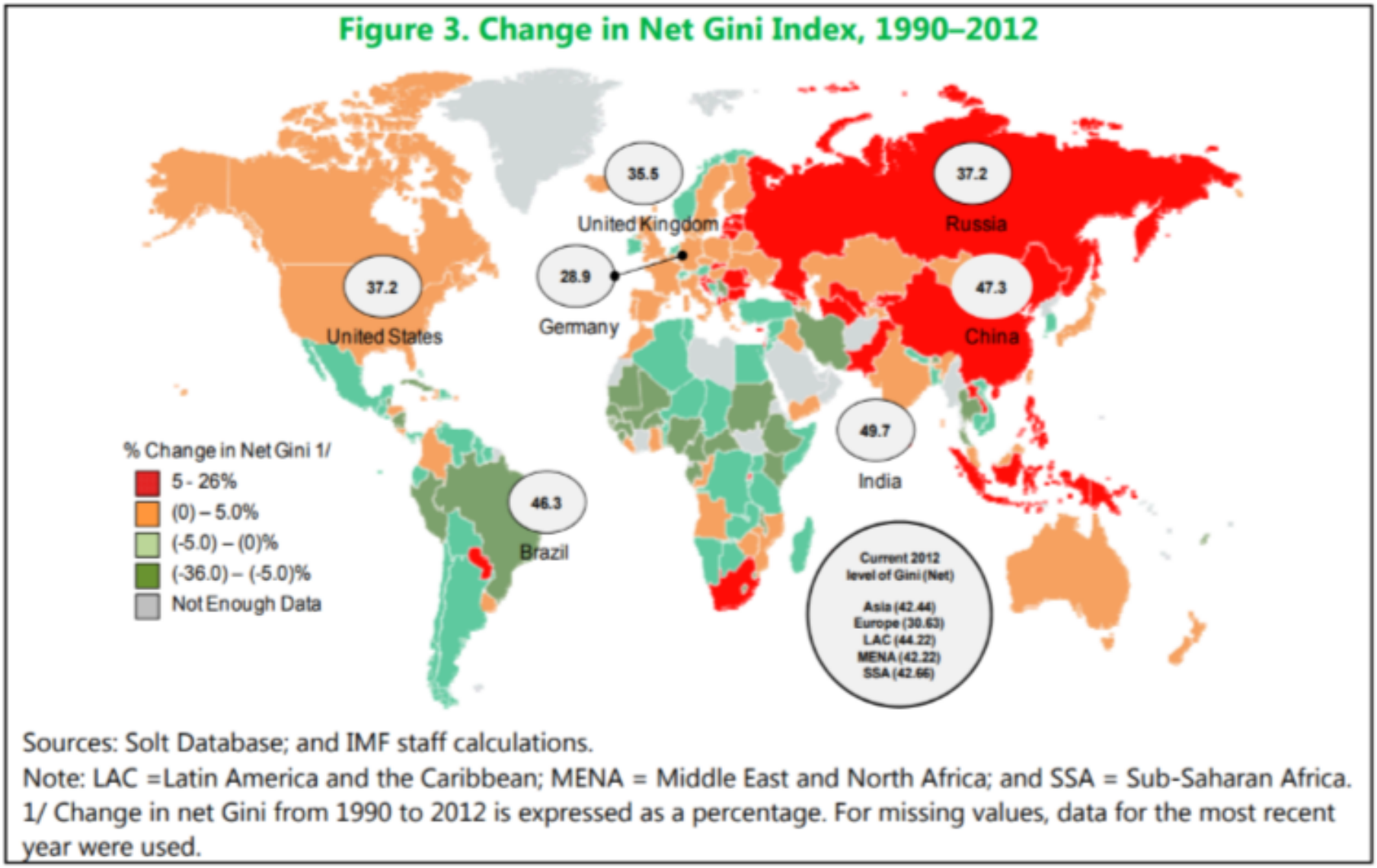
Satellite images of South Asia by night



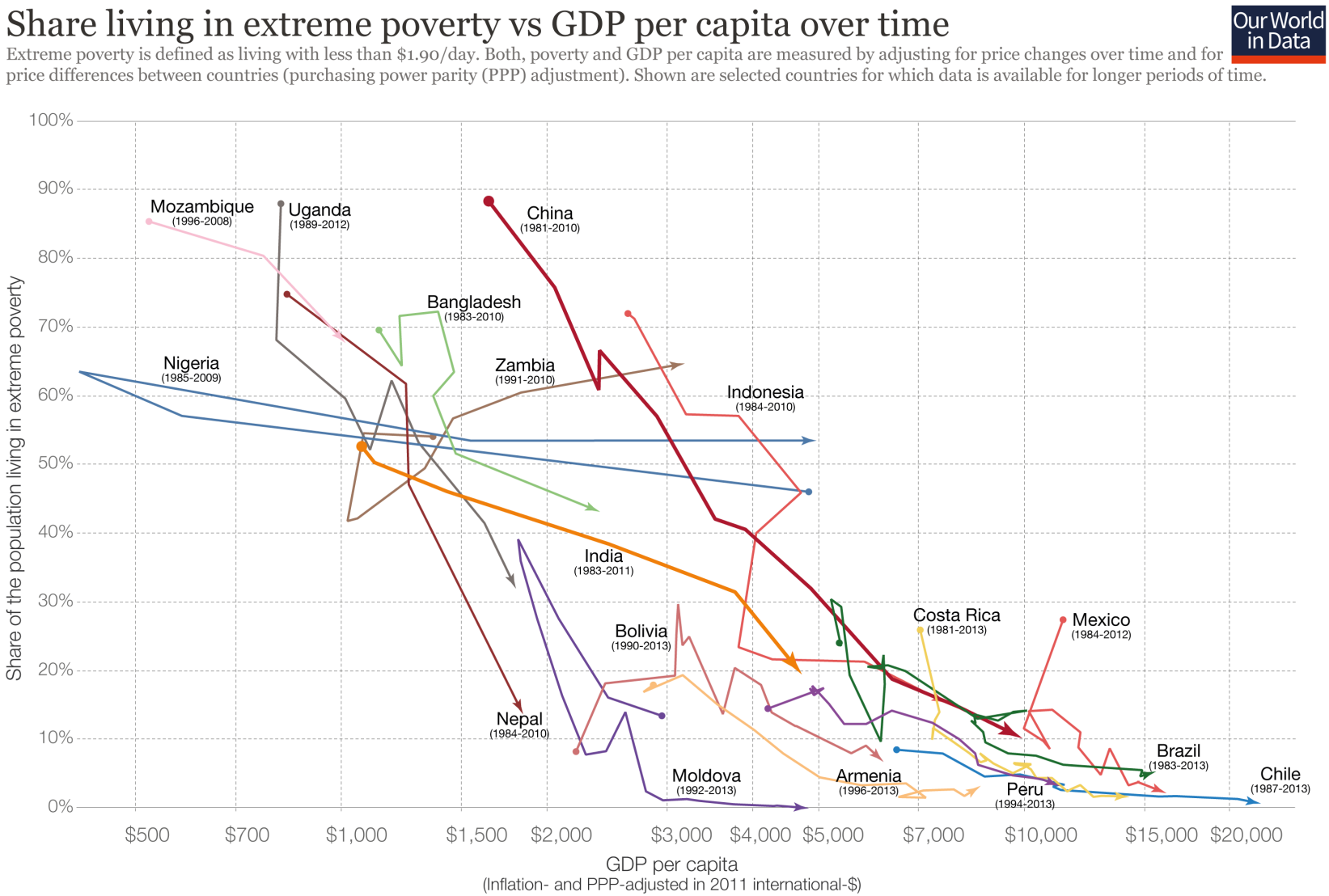
South Asia in 1994 South Asia in 2010

Images are taken from Maxim Pinkovskiy and Xavier Sala-i-Martin (2016) – *Lights, Camera … Income! Illuminating the National Accounts-Household Surveys Debate*. The Quarterly Journal of Economics

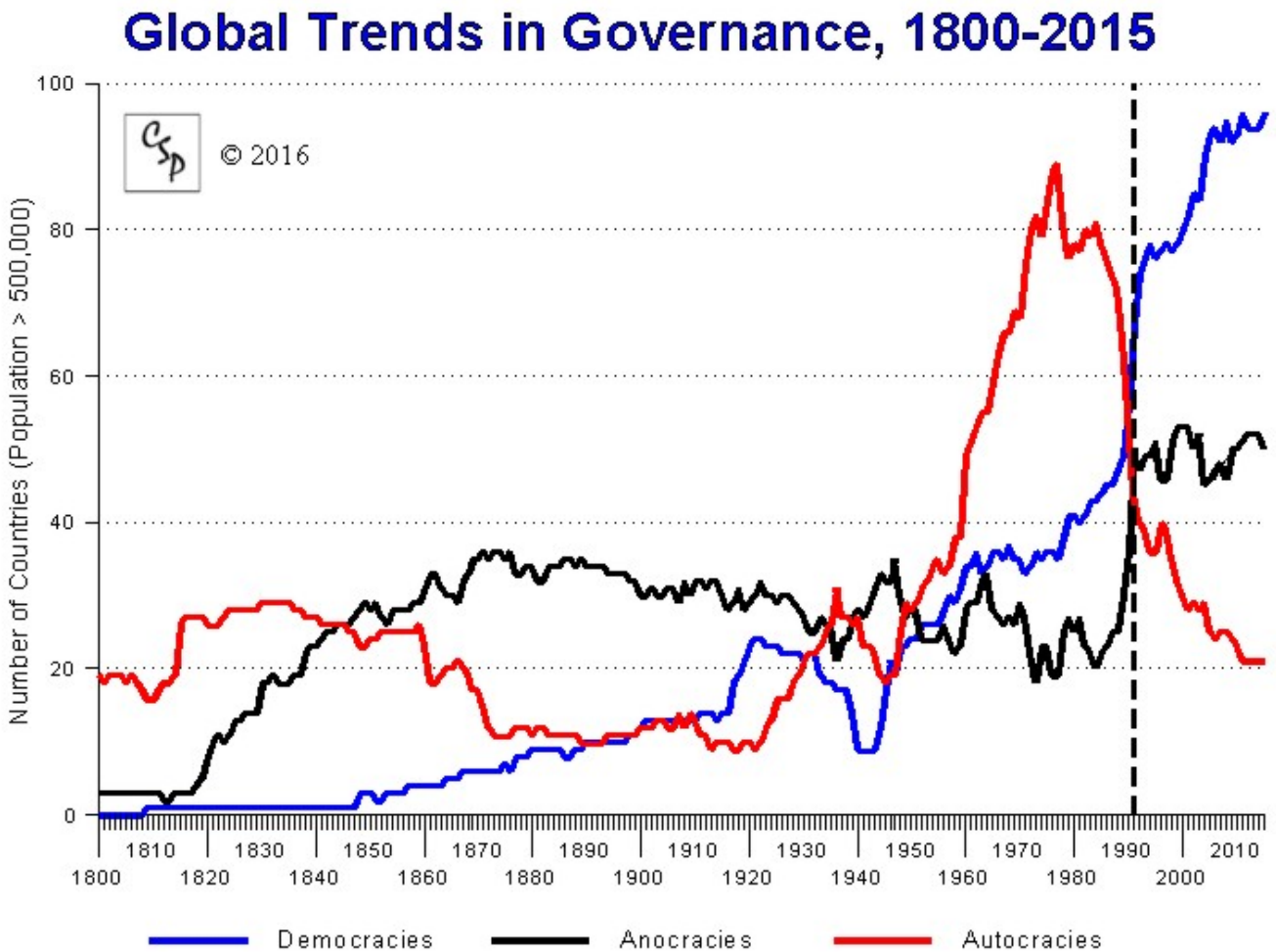
The Stylized Facts of Development



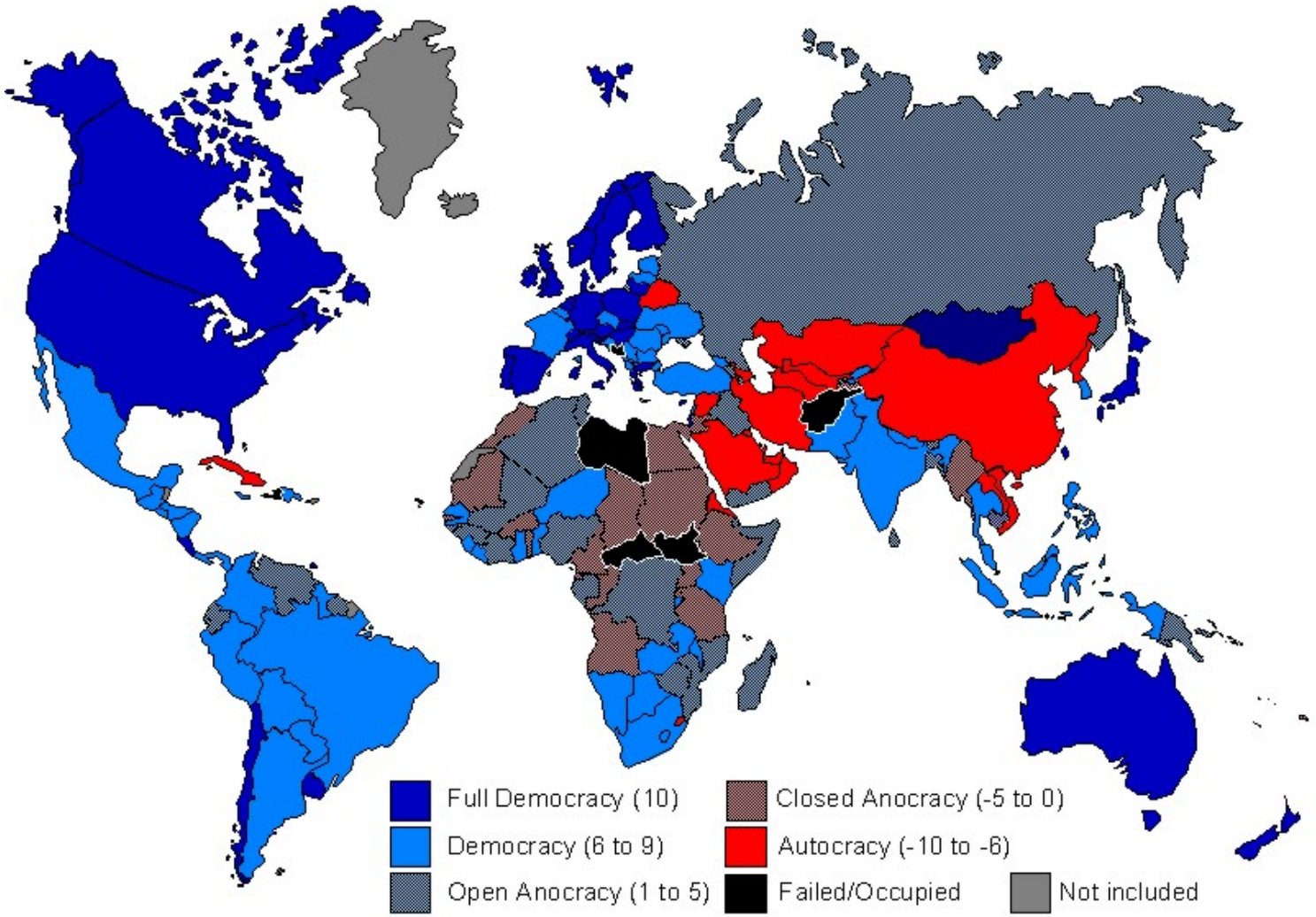
The Stylized Facts of Development



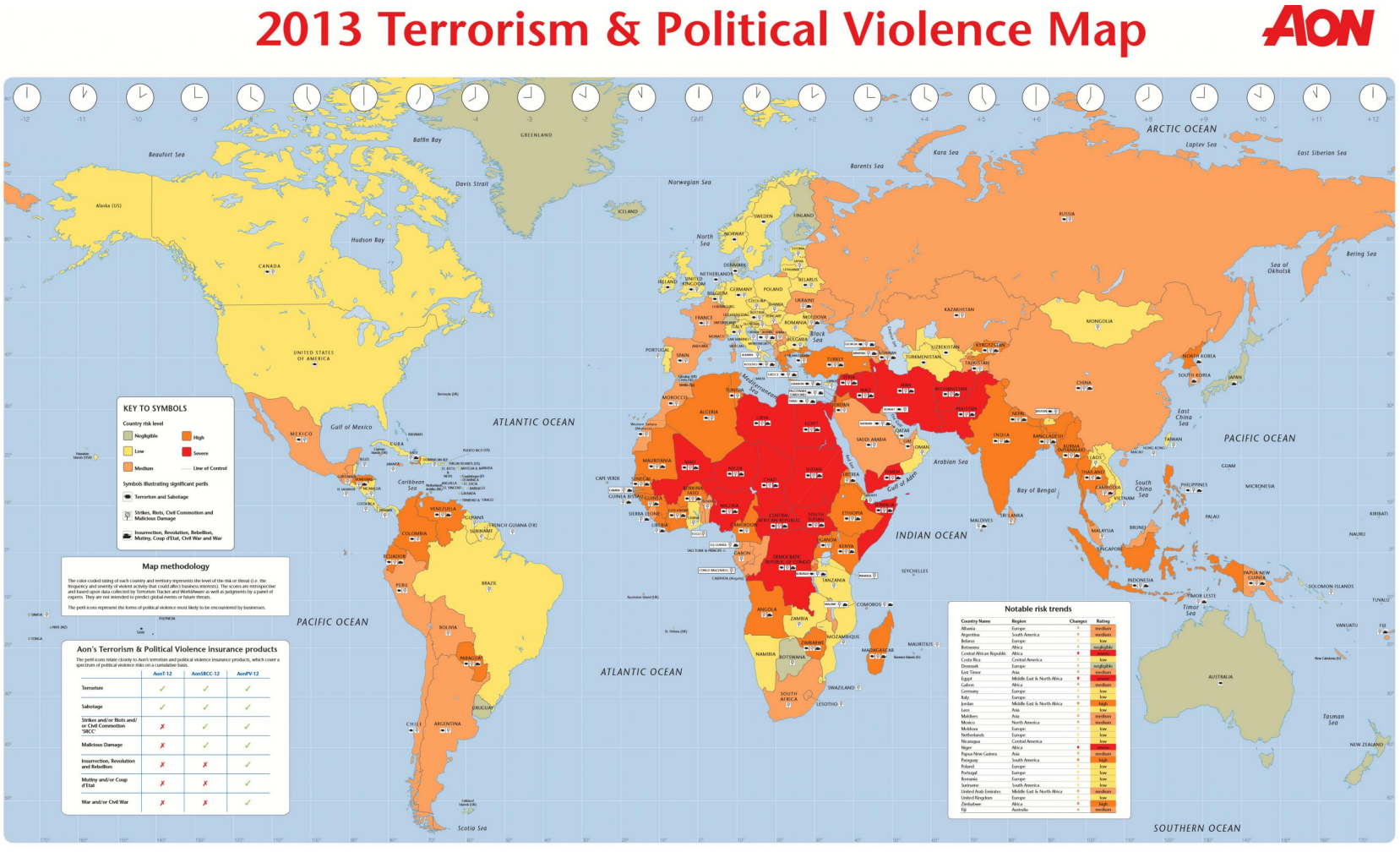
The Stylized Facts of Development



The Stylized Facts of Development



The Stylized Facts of Development



The Stylized Facts of Development

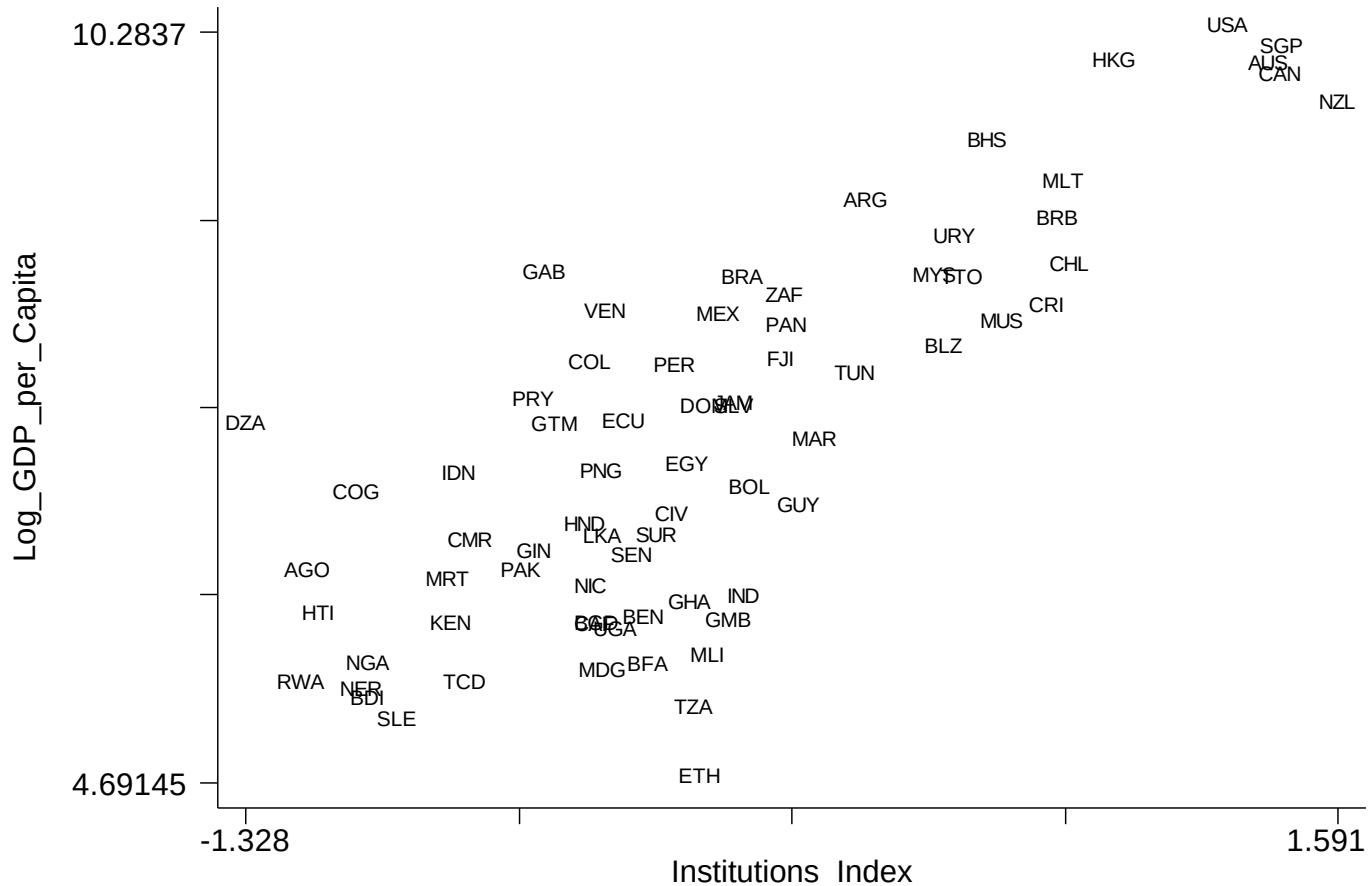
Selected facts to be explained:

- Number of poor starts falling around 1970
- Poverty clustered in Africa and South Asia
- Poverty affects the young more than the old, and women more than men
- Within-country variation is high
- Growth can be more or less poverty-reducing
- Democracy and violence is uneven, and not the same as poverty
- Full authoritarianism has declined since 1989, but anocracies are more common

Explaining Development

- These patterns are not random
- Geography matters, but the exceptions are many:
  - Haiti vs Dominican Republic
  - Botswana vs Nigeria
  - Venezuela (richest country in LA in 1970)
- Geography matters through its impact on **institutions and policies** (Easterly and Levine 2003, Hall and Jones 1999)
- History matters through its impact on **institutions and policies** (Acemoglu, Johnson and Robinson 2001)

Figure 3: Logarithm of GDP per Capita in 1995 vs. Institutions Index

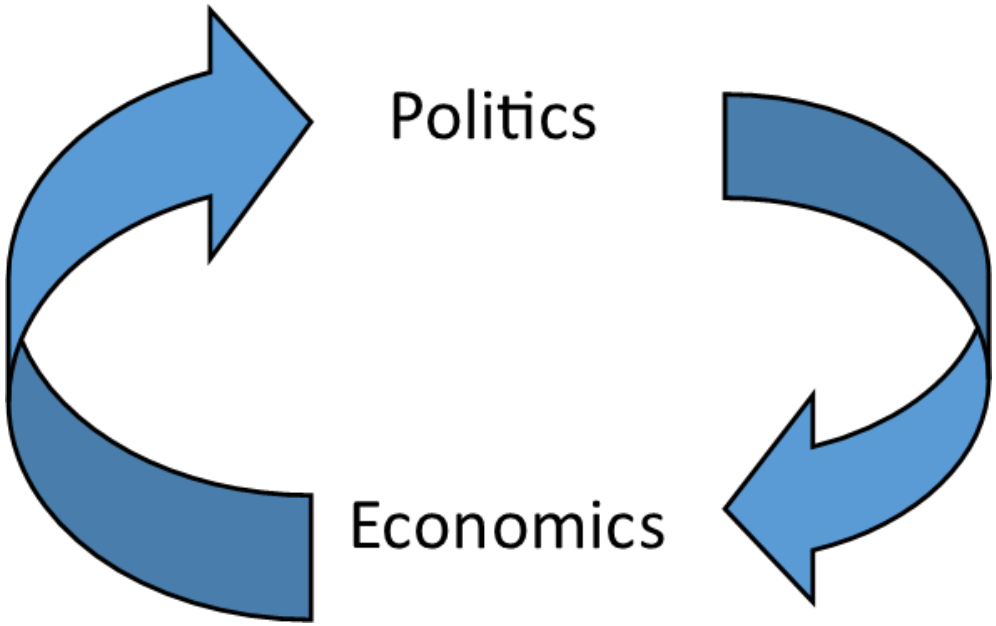


Political Economy

- **Political Economy of Development:** How politics influences economic policy and development outcomes

Political Economy

- **Political Economy of Development:** How politics influences economic policy and development outcomes
- But also: how economics shapes political choices



Explaining Development: Political Economy

- **Economics:** Focused on how resources are produced and distributed
  - Tells us how policy decisions affect development/poverty
  - Huge literature on **optimal policy choices**

1. But why are these policies chosen?
   - Government is not a benevolent dictator
   - Economic consequences influence policy choice
   - But knowing a policy benefits the poor isn't enough: how does that affect the policymaker?
   - Political effects are different to economic effects
2. And how are these policies actually implemented?
   - No magic wand; politicians and bureaucrats can distort implementation
   - Citizens can also resist and react

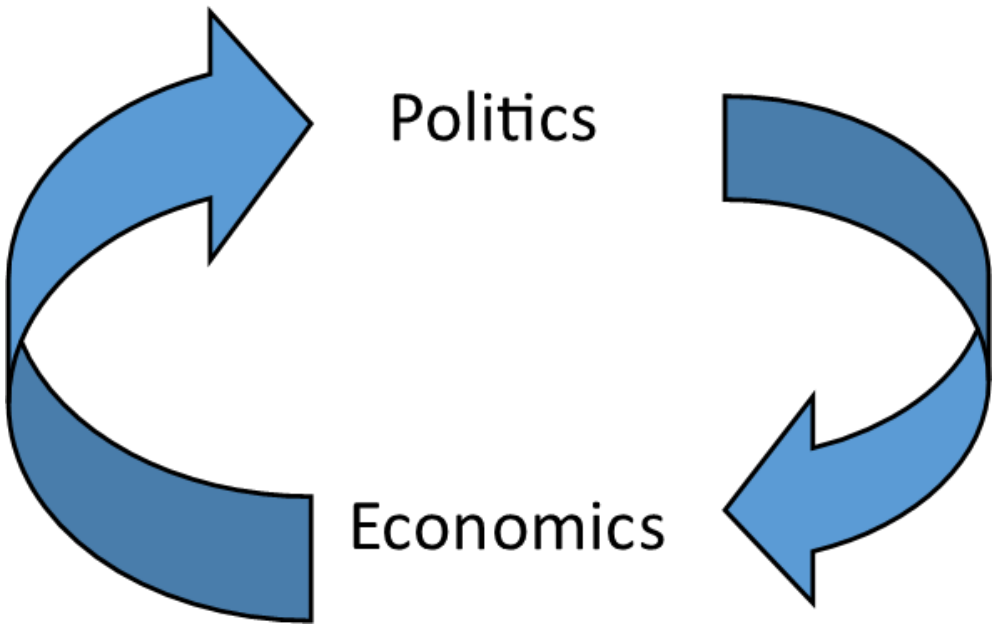
Political Economy

- **Politics:** How societies make decisions over 'who gets what'
- Those decisions are **described in institutions** - rules of who gets what
  - eg. tax rates, education subsidies, monetary policy
- Those decisions are also **altered by institutions** - rules of who gets to decide
  - Regime type, voting rights, minority protections, judicial independence
- Decisions that affect others are acts of **political power**

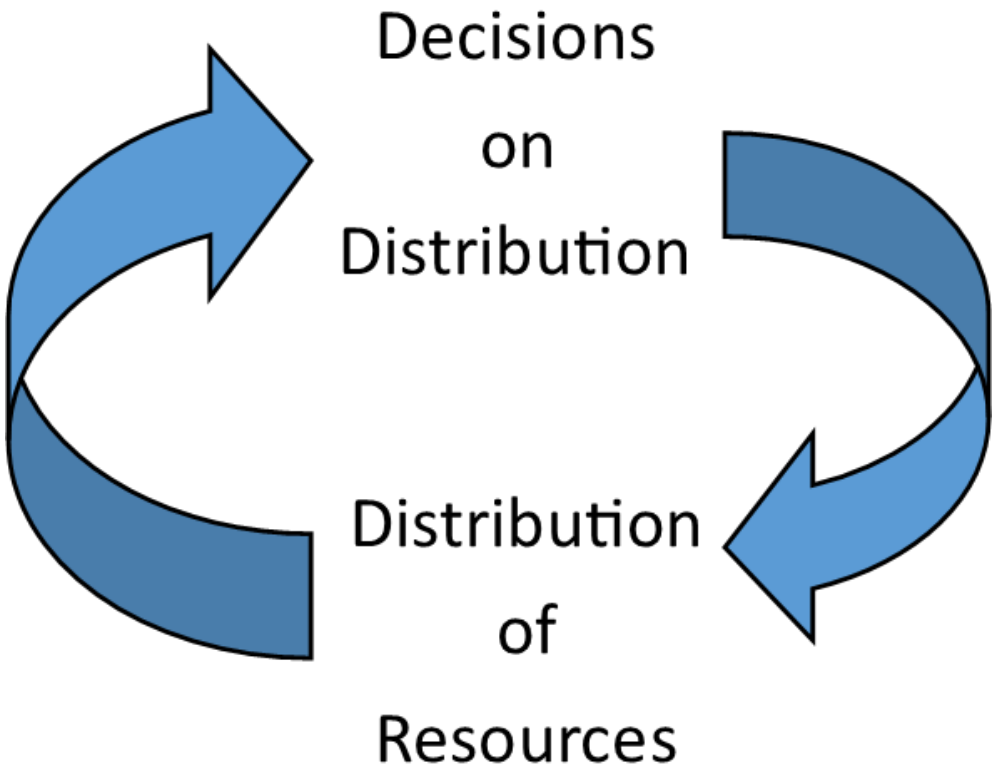
Political Economy

- **Politics:** Institutions are not enough
  - Institutional rules create opportunities, but taking advantange requires:
    - Agency
    - Collective Action
    - Resources
    - Information
  - Rules are never complete or fully enforced
  - Some actions can ignore rules - violence
  - Rules can be changed by mobilization
- So there is huge space for politics to occur ‘within’ or ’around’ institutions

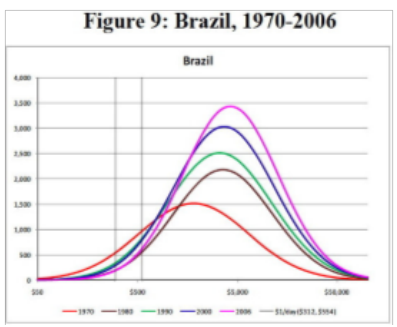
Political Economy



Political Economy



Political Economy



Political Economy

- This course is structured around three key political economy processes:
  1. How institutions affect development outcomes
  2. How people organize and exercise political power *within* institutions
  3. How institutions change

Political Economy

- Consider an example society:
  - One farmer, one factory worker and one tax collector
  - Farmers and tax collectors want high taxes, workers low taxes
- How do institutions affect development outcomes?
  - How does a dictatorship of factory workers affect development?
  - How does a dictatorship of farmers affect development?
  - How does a democracy affect development?

Political Economy

- How do people organize within institutions?
  - Will the tax collector enforce the tax policy?
  - Will the factory worker actually pay their taxes?

Political Economy

- How do institutions change?
  - A new policy idea becomes available to train the tax collector as a factory worker
  - Under a democracy, who will vote for this change?

Causal Inference

- Try defining your own question for the Political Economy of Development
- After, we'll try and turn it into a causal inference question

Causal Inference

- How do we learn about the political economy of development?

	Freedom of Information Law	Falling Poverty
A	0	0
B	0	0
C	0	0
D	0	0
E	1	1
F	1	1
G	1	1
H	1	1
I	0	?
J	1	?

Causal Inference

- **Correlation is not Causation**
- Why is causation so important?
  - Because development is *extremely complex*
- Description is helpful, but even just to inform policy we need to understand the full process
  - Changing a correlated variable may not produce any results
  - Lucas Critique: Observed relationships in data might not hold if you start changing policy because policy might affect the nature of the relationship
    - eg. The data shows no-one lies on their tax forms.
    - So why bother with tax checks, let's save some money, right?
    - But reducing checks reduces the chance of getting caught
    - And might lead to widespread lying

Causal Inference

- So we need to learn about the **causal mechanisms** that drive behaviour and shape outcomes
- The problem is not data *quality*, but how the data were generated
- We need data generated in ways that reveal the causal mechanism - what would happen if we changed a variable, keeping everything else the same

Causal Inference

- So the type of questions we are asking are NOT "What caused Y?"
  - eg. Why did the United States grow faster than Bolivia in the twentieth century?
- But "Does X affect Y?"
  - eg. Did the more permanent colonial settlement of the United States compared to Bolivia affect their subsequent growth rates?
- These are called “Effects of Causes” questions (not “Causes of Effects” questions)

Causal Inference

- A focus on a single explanatory variable X requires us to clearly define this 'treatment'
- AND to clearly define a control
  - What is the opposite of investing $1bn in education?
  - No investment, or investing it elsewhere?
- Define treatment:

$$D_i = \begin{cases} 1, \text{ if treated} \\ 0, \text{ if not treated} \end{cases}$$

Causal Inference

- Defining our outcome is also crucial:
  - Can we measure our outcome of interest?
  - Is that outcome the end of the causal chain?
  - Tempting to look at many outcomes, but the risk of cherry-picking
    - All outcomes are probabilistic
    - If we study 20 outcomes, on average one will show a significant effect even with no real causal effect

Causal Inference

- Learning about causal effects requires us to specify the 'unit' - what is being affected?
- Countries? Political Parties? Individuals?
- eg. How does segregation affect attitudes to redistribution?
  - Treatment at the community/societal level
  - Outcome at the individual level
  - Measurement needed at the individual level
- Units are **time-specific**: the same person 10 minutes later is a different unit

Causal Inference

- We want to know how some variable affects another variable
- eg. how a proportional representation electoral system affects investment in education
  - The **unit** here is any political system where both electoral system and education can vary independently of other units, i.e. countries
  - The **treatment** is a change to a PR electoral system (vs FPTP)
  - The **outcome** is the level of (public?) investment in education

Causal Inference

- Causality is complex, eg. for $X \rightarrow Y$:
  1. Many factors influence a single outcome ($X1, X2 \rightarrow Y$)
     - Parliamentarism also influences investment in education
  2. Equifinality: Many routes to the same outcome ($X1 + X2$ or $X3 + x4 \rightarrow Y$)
     - Ghana and Iceland spend the same on education, but in very different ways
  3. Reverse causation ($Y \rightarrow X$)
     - A highly educated population might prefer a PR system
  4. Non-linear impact of one variable on another ($X \Rightarrow Y$)
     - A mixed electoral system may have no effect, but a full PR system might lead to a big jump in investment
  5. General equilibrium effects - treatment affects many other variables ($X \rightarrow Y1, Y2 \rightarrow Y1$)
     - Public investment in education rises, but private investment falls by the same amount

Causal Inference

6. Context matters ($X|Z \rightarrow Y$)
   - PR has a different effect in British vs French legacy education systems
7. Treatments cannot be replicated ($X1 \rightarrow Y1, X2 \rightarrow Y2$)
   - Some countries apply open list PR, others closed list etc.
8. Spillovers between units ($X_T \rightarrow X_C \rightarrow Y$)
   - When New Zealand switched to PR, Australia was a natural comparator, but to compete for students, Australia also raised education investment
9. Learning, demonstration effects and history matter ($X_{t=1} \rightarrow Y1, X_{t=2} \rightarrow Y2$)
   - New Zealand adopted PR *because* it saw that education improved in Japan
10. Social complications eg. emotion, irrationality, chaos theory ($X \rightarrow Y1, X \rightarrow Y2$)
    - New Zealand introduced PR because of an off-hand remark by one person in a campaign

Causal Inference

- So we need a precise framework to analyze causation
- The causal effect of treatment is how the unit's outcome differs when it is treated and not treated
- These are the **potential outcomes** for unit i:

$$Y_{Di} = \begin{cases} Y_{1i} \text{ Potential Outcome if unit i treated} \\ Y_{0i} \text{ Potential Outcome if unit i not treated} \end{cases}$$

- Treatment Effect $= Y_{1i} - Y_{0i}$

Causal Inference

Potential Outcomes Example

	Investment in Education if PR system	Investment in Education if FPTP system	
	Y_1	Y_0	Treatment Effect
Brasil	8	4	4
Argentina	10	7	3
Bolivia	2	4	-2
Colombia	11	11	0
Peru	6	2	4

Causal Inference

- **The Fundamental Problem of Causal Inference**
  - No units can receive **both** treatment and control
  - So we can never observe both Y_1 and Y_0 for the same unit

Causal Inference

Potential Outcomes Example

	PR System?	Investment in Education if PR system	Investment in Education if FPTP system	
	D_i	Y_1	Y_0	Treatment Effect
Brasil	1	8	?	?
Argentina	1	10	?	?
Bolivia	0	?	4	?
Colombia	0	?	11	?
Peru	0	?	2	?

Causal Inference

- We can't even look at the change in countries that switch to a PR system
  - What if **all** countries had started to invest more in education at the same time, for different reasons?
  - The potential outcome for Country X in time 1 is different to at time 2
- So we need to consider the **counterfactual** - what would have happened if the country had **not** switched to a PR system?
- So we can only estimate the effect by comparing **across** units
- That is why we are doing causal **inference**, not causal proof

Causal Inference

- To compare across units we need counterfactuals: **control** units that do not receive treatment
- Control units can never be perfect substitutes
- Causal Inference is all about identifying a **plausible counterfactual**
  - Plausible means that the potential outcomes of the control unit are the same as those of the treated unit

Causal Inference

- The comparability of treatment and control units depends on how they got to be treated
  - On the **treatement assignment mechanism**
- If we 'treated' an outlier like Búzios in Rio, could we find a comparable control unit?
- Comparisons are easier where the **treatment assignment mechanism is independent of potential outcomes**
  - This makes it more likely that potential outcomes are 'balanced' and comparable

Causal Inference

- The rest of the course is mostly about the types of treatment assignment mechanisms that **avoid these biases** and provide plausible counterfactuals

Causal Inference

1. **Controlled Experiments** where we **control** the treatment assignment
   - Field Experiments
   - Survey Experiments
   - Lab Experiments

Causal Inference

2. **Natural Experiments** where the assignment mechanism creates balanced potential outcomes
   - Randomized natural experiments
   - Regression Discontinuities
   - Instrumental Variables

Causal Inference

3. **Observable Studies:** What if no suitable treatment assignments are available?
   - No historical examples of natural experiments
   - Not feasible or ethical to run a field experiment

- Remember the purpose of using these specific treatment assignment mechanisms is to achieve **comparable potential outcomes**
- One alternative way of making potential outcomes comparable is to **selectively use Observable Data**
  - Difference-in-Differences
  - Controlling for confouding variables
  - Matching

Causal Inference

Analysis Types and Assumptions

Week	Assumption:	Researcher Controls Treatment Assignment?	Treatment Assignment Independent of Potential Outcomes	SUTVA	Additional Assumptions
	Controlled Experiments				
1	Field Experiments	✓	✓	✓	
2	Survey and Lab Experiments	✓	✓	✓	Controlled Environment for treatment exposure
	Natural Experiments				
3	Randomized Natural Experiments	X	✓	✓	
4	Instrumental Variables	X	✓	✓	First stage and Exclusion Restriction (Instrument explains treatment but not outcome)
5	Regression Discontinuity	X	✓	✓	Continuity of covariates; No manipulation; No compounding discontinuities
	Observational Studies				
6	Difference-in-Differences	X	X	✓	No Time-varying confounders; Parallel Trends
7	Controlling for Confounding	X	X	✓	Blocking all Back-door paths
8	Matching	X	X	✓	Overlap in sample characteristics

Causal Inference

4. **Small-N studies:** Some research questions have few units available

- How do we learn about the political economy of development with few units?
- We can at least avoid some key biases:
  - Comparative Case Studies
  - Process Tracing

Causal Inference

- But **how much** can we learn from a causal analysis?
- Is this an accurate representation of what would happen in the real-world?
  - What was the policy problem (/academic question) you were trying to solve?
  - What details differ? Eg. context of how treatment was applied
- Generalizability to other units (External validity)
  - Would the same thing happen in another country? Next year?
  - Look out for variation in treatment, context, spillovers, learning etc.
- Any generalization requires assumptions

Causal Inference

- We will try to identify abstract, portable processes
  - **Causal Mechanisms**
- **Portable:** If the weather affects election turnout ONLY in Acre, is that a useful causal mechanism?
- **Abstract:** If unions are good at mobilizing support, but so are churches, the mechanism is collective action, not union organization
- We still need to define the **scope conditions** in which we think this causal mechanism will operate as expected

Causal Inference

- Examples of Causal Mechanisms:
  - Citizens
    - Electoral Accountability
    - Client Power
    - Collective Action
    - Social Trust/Sanctioning
    - Wealth Effects
  - Elites
    - Violence/Coercion
    - Brokerage/Patronage
    - Persuasion/Framing
    - Incumbency Power
  - Institutions
    - Power Devolution/Median Voter
    - Network Effects
    - Evolutionary Selection
    - Conversion/Layering/Drift/Replacement

Causal Inference

- Examples of Causal Mechanisms:
  - Citizens
    - Electoral Accountability - Class 5
    - Client Power - Class 6
    - Collective Action - Class 11
    - Social Trust/Sanctioning - Class 4
    - Wealth Effects
  - Elites
    - Violence/Coercion - Class 8
    - Brokerage/Patronage - Class 9
    - Persuasion/Framing
    - Incumbency Power - Class 7
  - Institutions
    - Power Devolution/Median Voter - Class 3
    - Network Effects
    - Evolutionary Selection
    - Conversion/Layering/Drift/Replacement - Class 12

Causal Inference

- Let's re-write your Political Economy Question as a Causal Inference Question

Reproducible Research

- So we have a substantive topic and a method. Now we need to implement it.
- The big problem: Give 5 researchers the same data and the same method and you'll get 5 different answers
- Replicating someone else's results is a minimum requirement, but it's hard
  - Manual data processing
  - No documentation of data processing
  - Errors unseen
  - Updates not consistent
  - Copy-paste errors
- Our research must be **reproducible**
  - Always generate the same results
  - Easily diagnose errors
  - Easily collaborate

Reproducible Research: R Markdown

- Principles of Reproducible Research
  1. Never touch the raw data
  2. Write code in a script
  3. Directly produce output documentation
  4. Every result comes from your code
  5. Comment and explain your code
  6. Manipulate data using clear rules, not individual items
  7. No cut-and-paste (more than twice)

Reproducible Research

- The tools we will use:
  - R (+Rstudio) - For data analysis
  - R markdown - simple way of writing reports
  - Latex - In the background, helps us create PDFs

Reproducible Research

- Installation
  - https://cran.r-project.org/bin/windows/base/ (Windows)
  - https://www.rstudio.com/products/rstudio/download/download
  - http://www.xm1math.net/texmaker/download.html

Reproducible Research: R

- R Cheat Sheet
- R_example.R

Reproducible Research: R Markdown

- Rmarkdown Cheat Sheet
- R_markdown_example.rmd

Reproducible Research: R Markdown

- In-Class Task:
  - Open a new R markdown document
  - Conduct a super-quick analysis of the WorldPhones data
  - Write up the results of your analysis (two sentences)
  - Include at least one table and one chart
  - Output the document to PDF, HTML and DOC files
- Use the example as a guide if you get stuck