1

1

2

3

**1**

We can fall in love with **stories** we tell ourselves about the world; however, if these stories remain **untested assumptions** we have no idea whether they are **true or false**.

**1**

How can we link solid **causal theories** to **real-world evidence**?

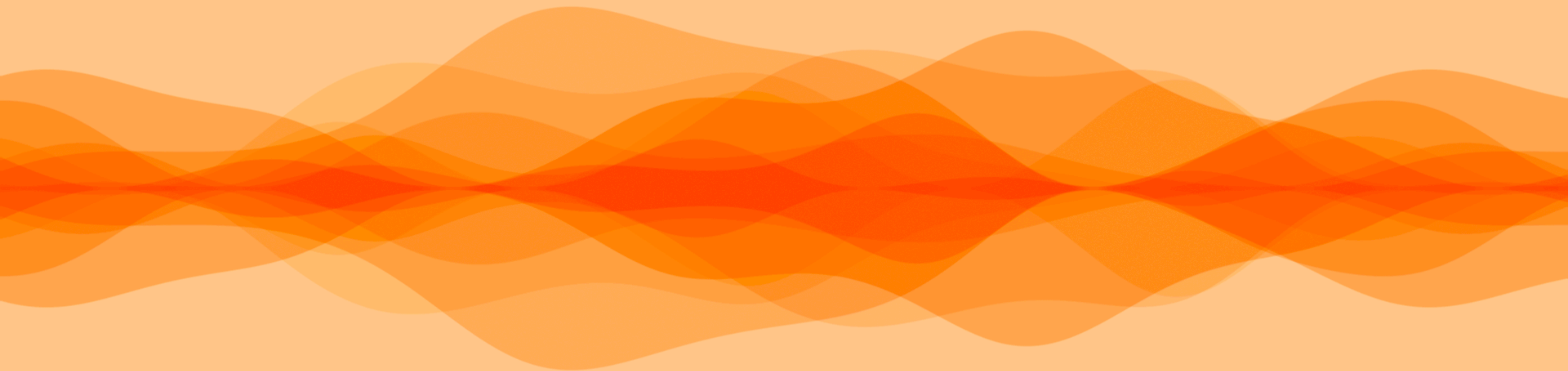How can we be sure this evidence has **measurement validity**?

**1**

Most people use real-world data without thinking about **how they are generated** and **whether they capture what they think they do**.
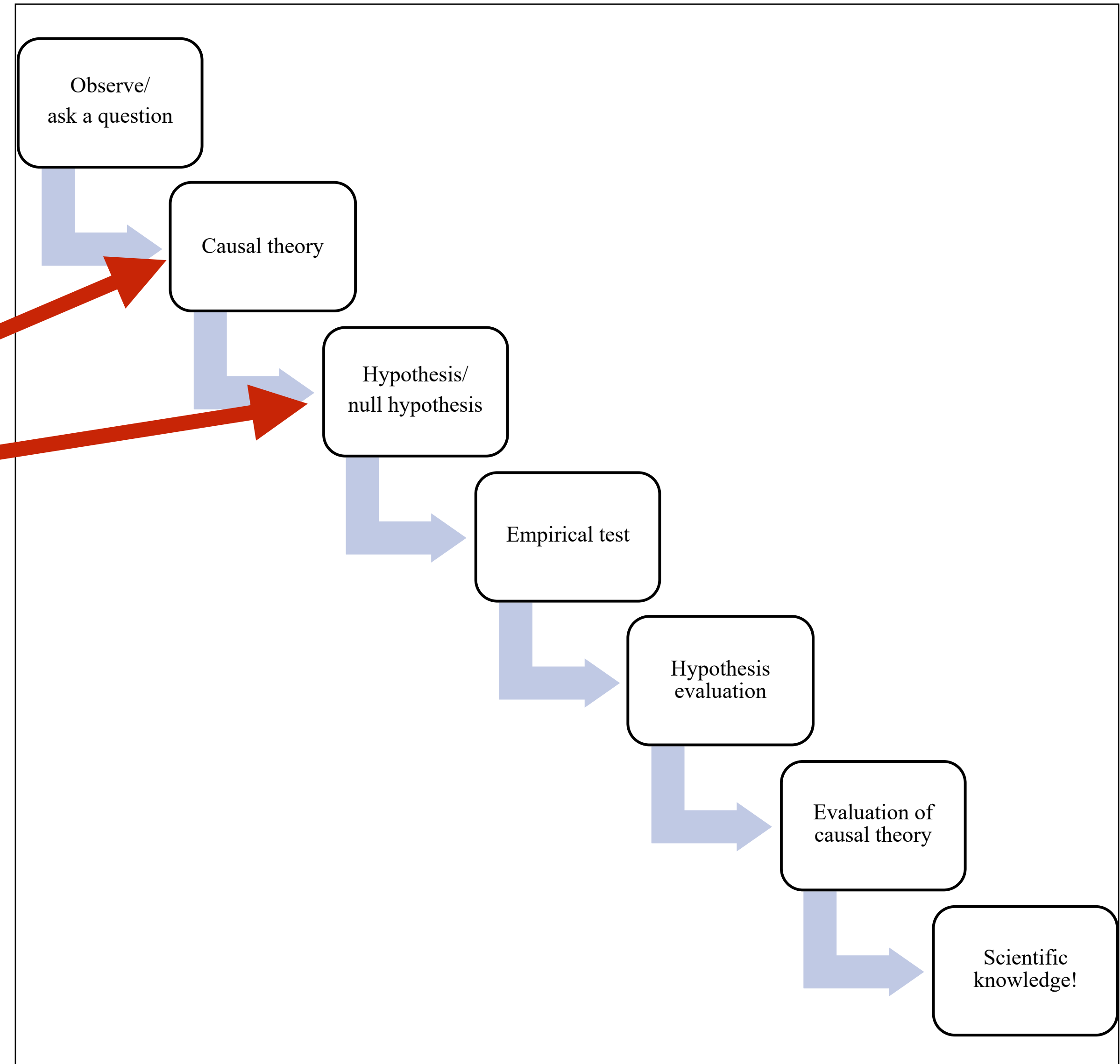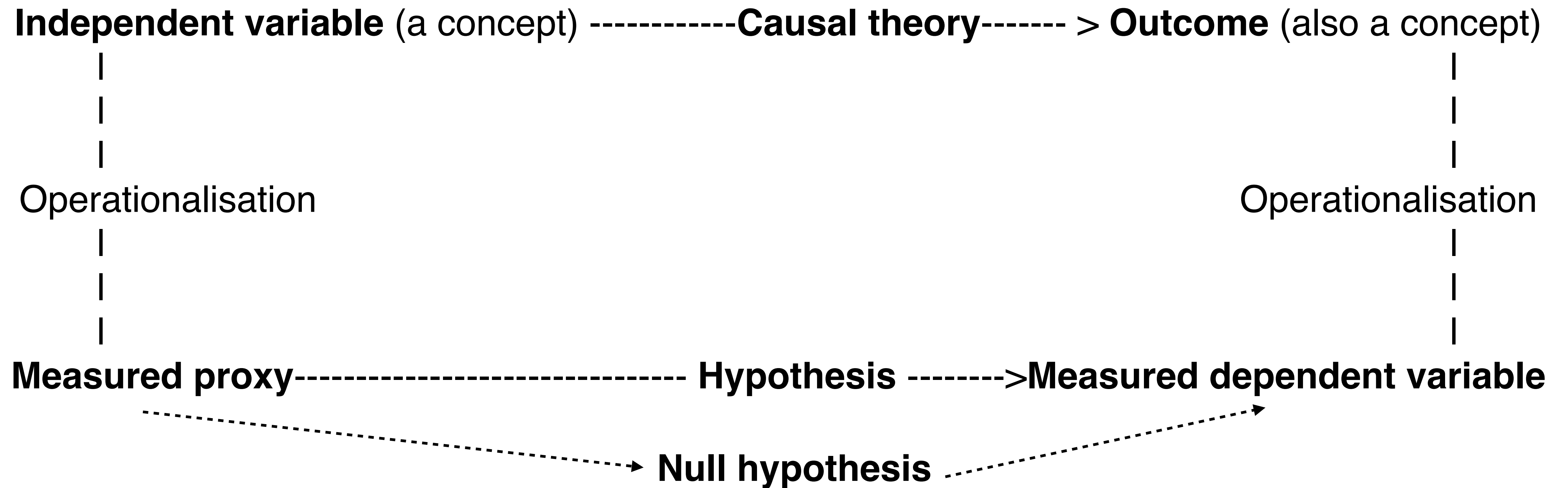
2

Observe/
ask a question

Causal theory

Hypothesis/
null hypothesis

Empirical test

Hypothesis
evaluation

Evaluation of
causal theory

Scientific
knowledge!

What is the difference?

**2**

**Independent variable** (a concept) ------------**Causal theory**------- > **Outcome** (also a concept)

Operationalisation

Operationalisation

**Measured proxy**-------------------------------- **Hypothesis** ------->**Measured dependent variable**

**Null hypothesis**

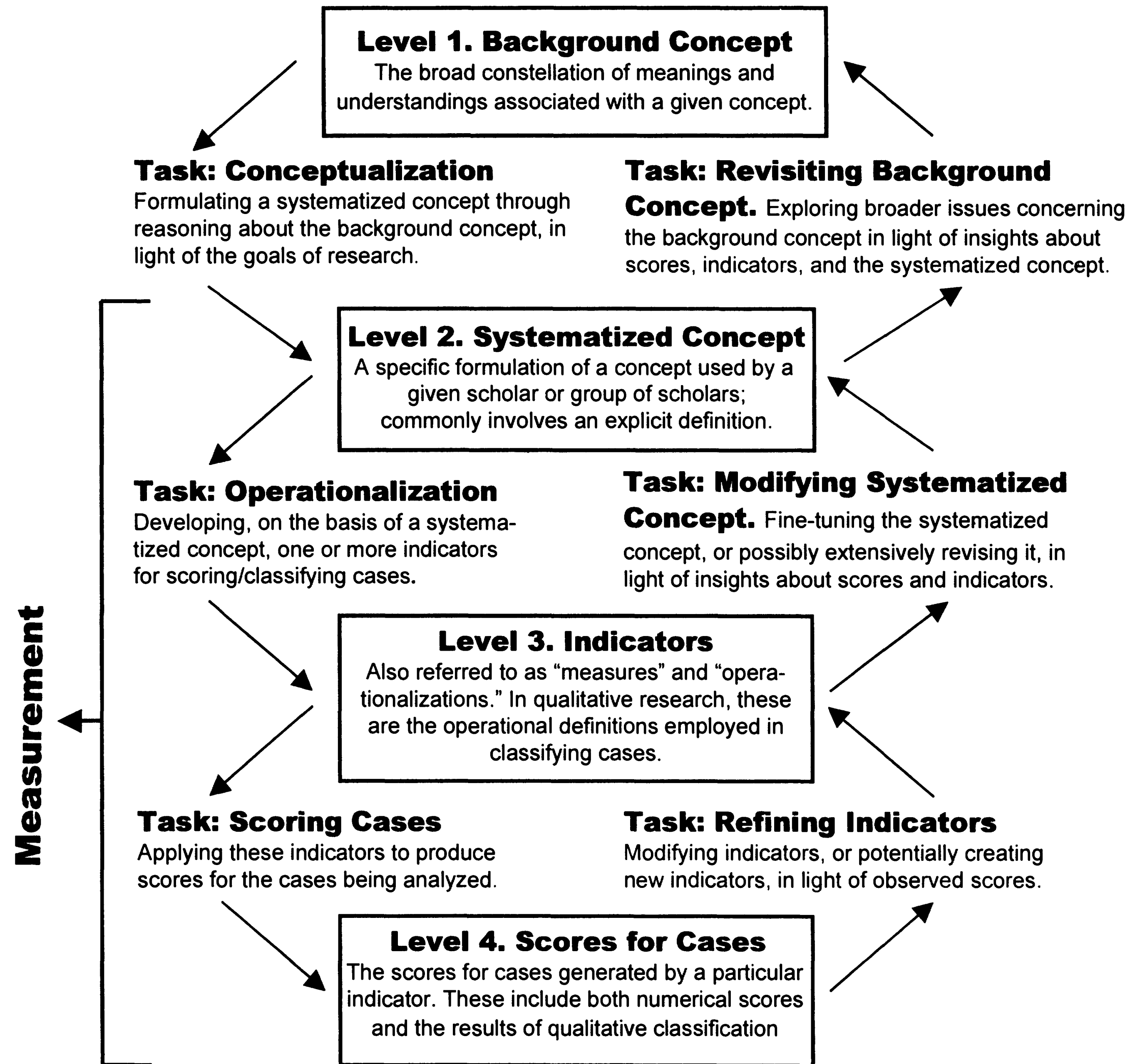Figure adapted from Kelstedt & Whitten (2018: 10).

**2**

Causal theories are relationships between **concepts**.

**Measurement** is trying to observe these concepts (or their proxies) in the real world.

**FIGURE 1. Conceptualization and Measurement: Levels and Tasks**

**Level 1. Background Concept**
The broad constellation of meanings and understandings associated with a given concept.

**Task: Conceptualization**
Formulating a systematized concept through reasoning about the background concept, in light of the goals of research.

**Task: Revisiting Background Concept.** Exploring broader issues concerning the background concept in light of insights about scores, indicators, and the systematized concept.

**Level 2. Systematized Concept**
A specific formulation of a concept used by a given scholar or group of scholars; commonly involves an explicit definition.

**Task: Operationalization**
Developing, on the basis of a systematized concept, one or more indicators for scoring/classifying cases.

**Task: Modifying Systematized Concept.** Fine-tuning the systematized concept, or possibly extensively revising it, in light of insights about scores and indicators.

**Level 3. Indicators**
Also referred to as "measures" and "operationalizations." In qualitative research, these are the operational definitions employed in classifying cases.

**Task: Scoring Cases**
Applying these indicators to produce scores for the cases being analyzed.

**Task: Refining Indicators**
Modifying indicators, or potentially creating new indicators, in light of observed scores.

**Level 4. Scores for Cases**
The scores for cases generated by a particular indicator. These include both numerical scores and the results of qualitative classification

**Measurement**

Source: Adcock & Collier (2001: 531).

12

**2**

**Step 1:** Conceptualise

**Step 2:** Operationalise

**Step 3:** Code

**1: Conceptualise** coffee shop characteristics
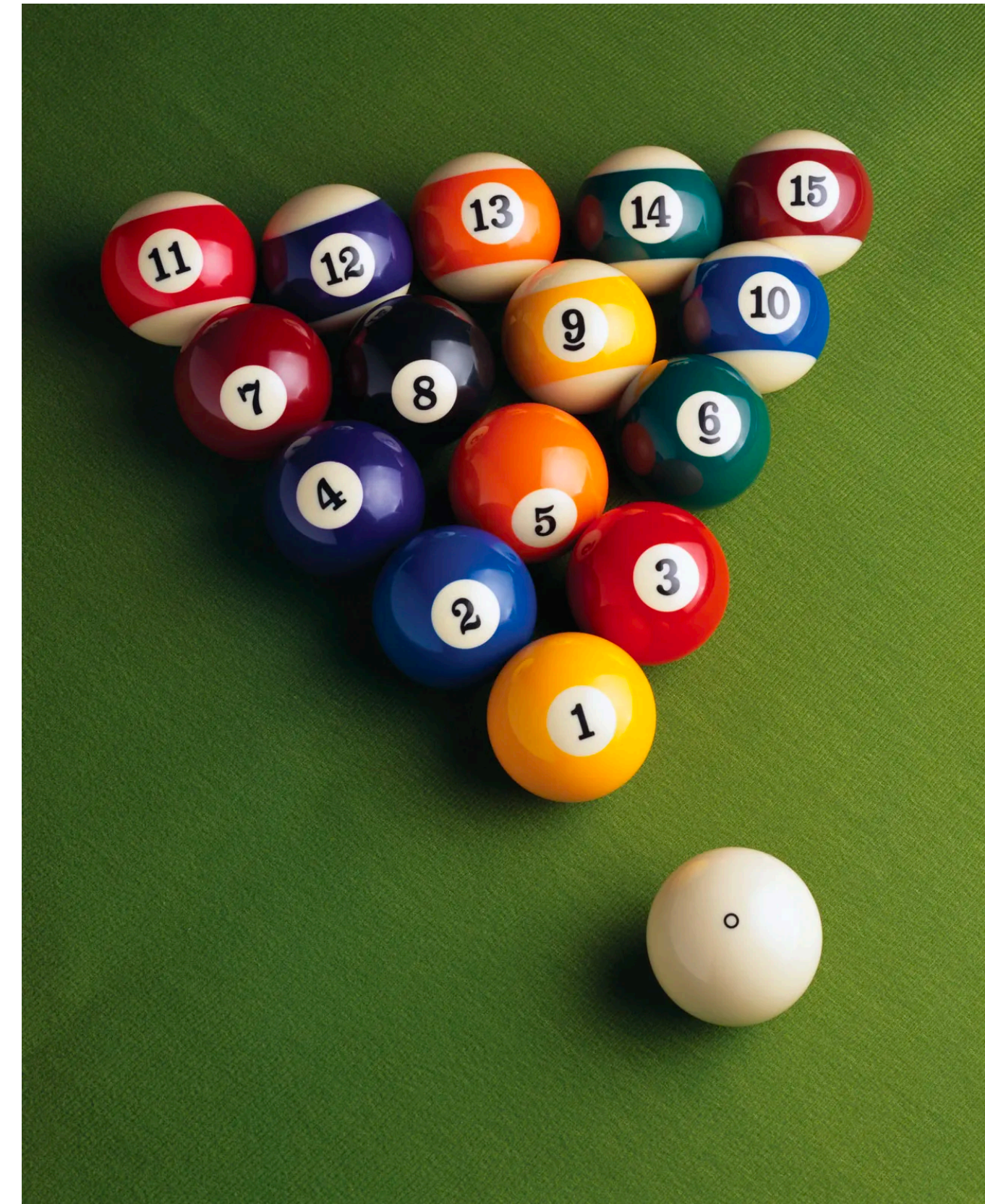
**2: Operationalise** these characteristics

**3: Code** as many coffee shops as possible
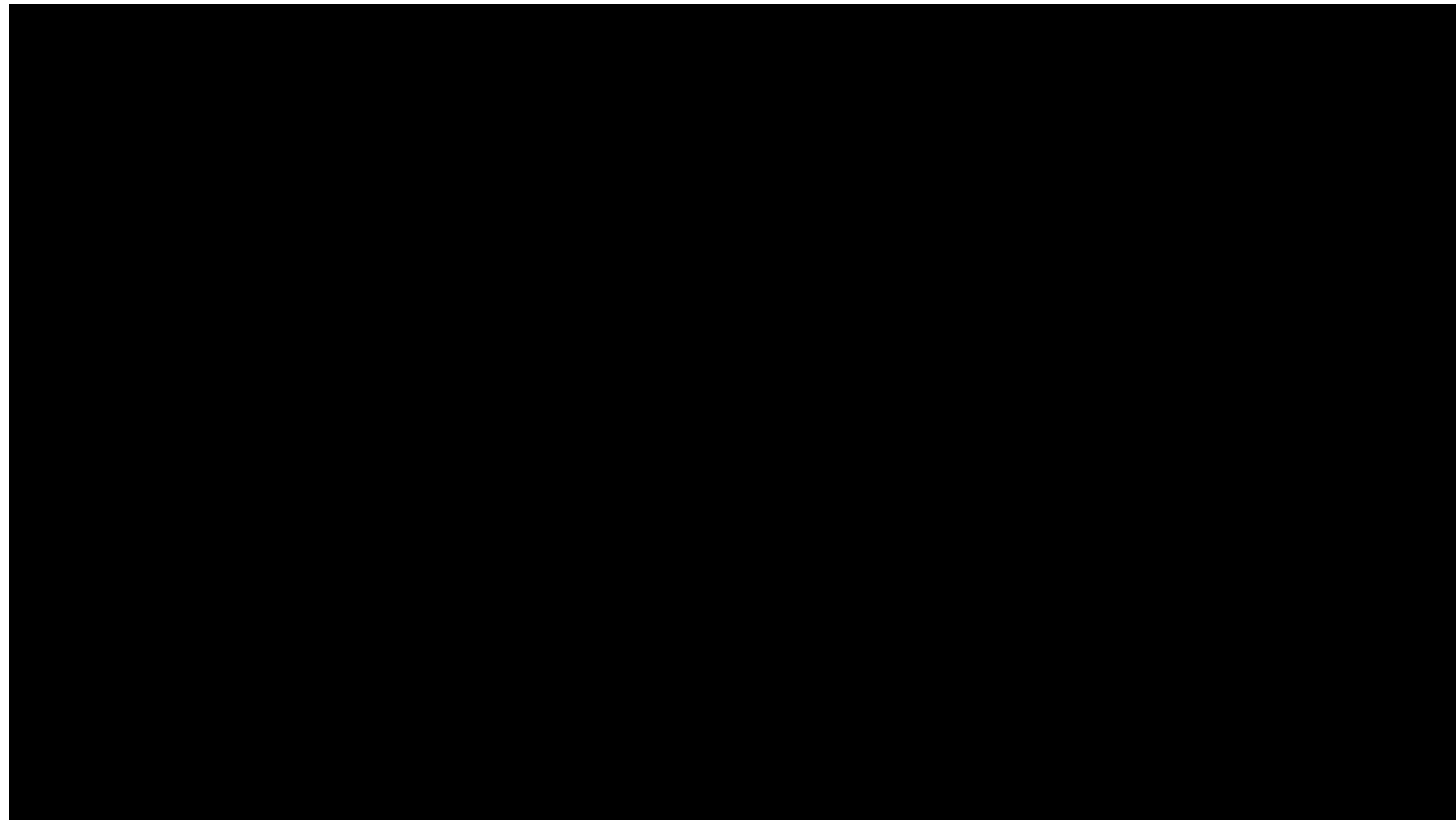
**Heterogenous** observations
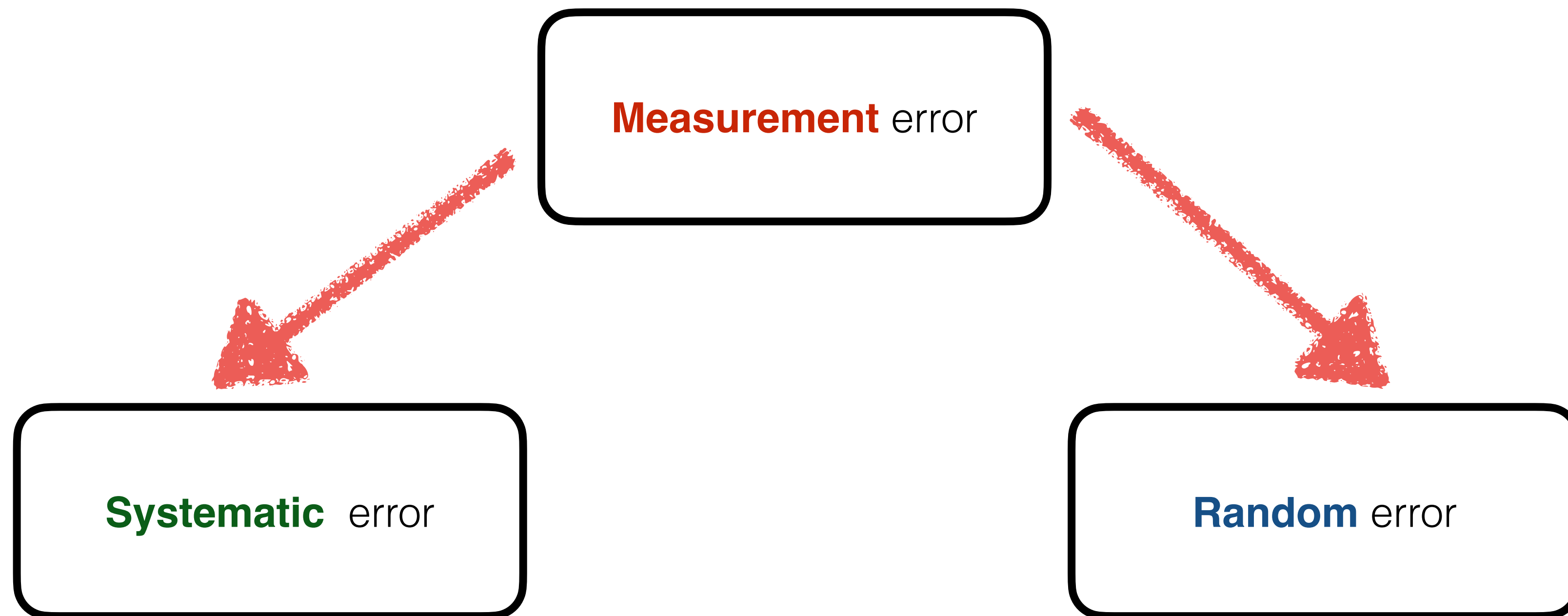


**Homogenous** observations

**2**

Any **gap** between
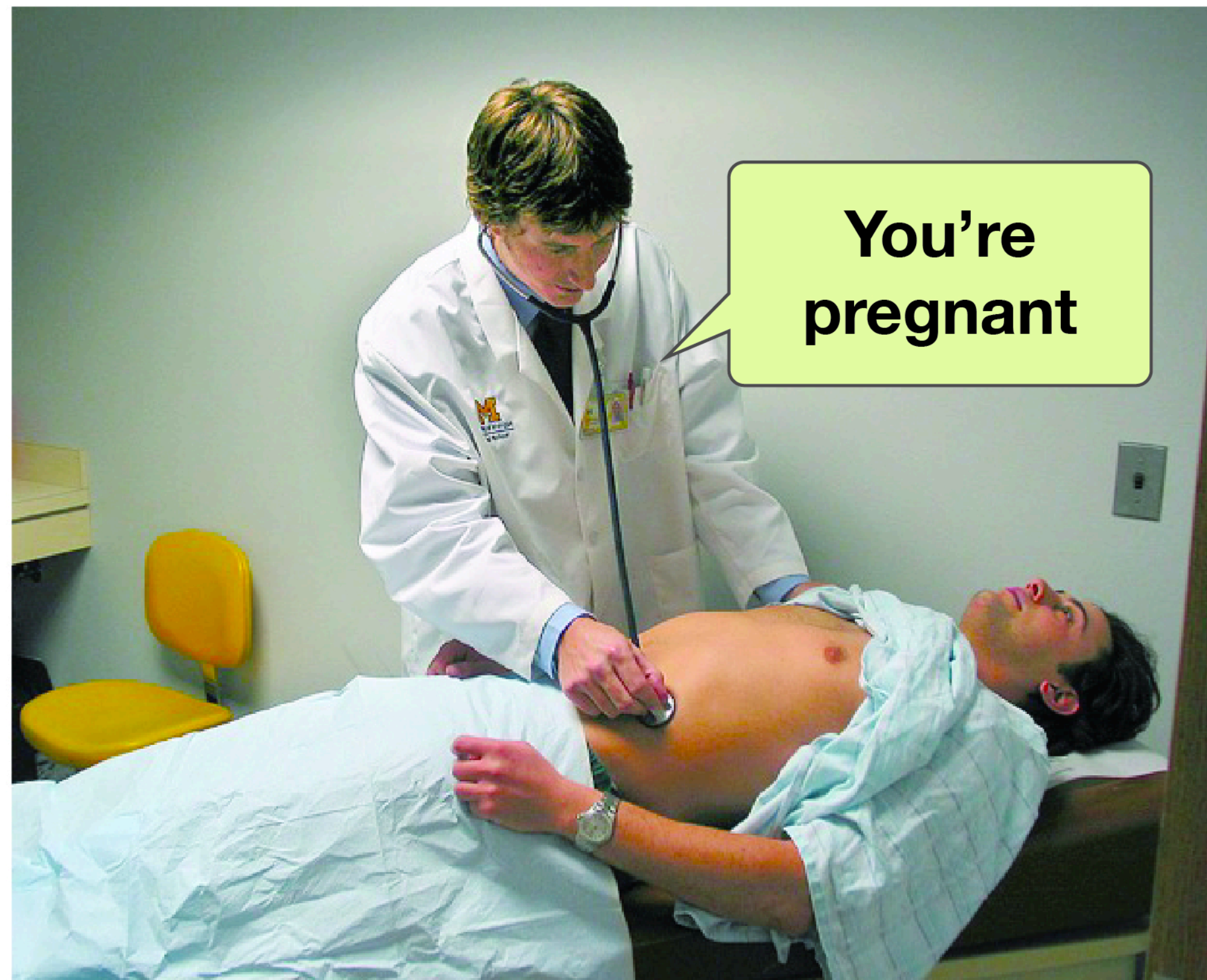theory and measurement here?

Source: https://youtu.be/KOO5S4vxi0o

Measurement error is the difference between the measured value and the true value of something.

**Measurement** error

**Systematic** error

**Random** error

# 2

# War Is in the Error Term
Erik Gartzke

The main theoretical task facing students of war is not to add to the already long list of arguments and conjectures but instead to take apart and reassemble these diverse arguments into a coherent theory fit for guiding empirical research.[1]

At least since Thucydides, students of international relations have sought rational explanations for the advent of war. Rationalist explanations assume purposive action; states are said to make reasoned decisions about the use of force. Although rationalist explanations have proven persuasive, durable, and offer the basis for cumulative theorizing, they also imply substantial limits on what we can know about war. I show that the most general rationalist explanation for war also dictates that the onset of war is theoretically indeterminate. We cannot predict in individual cases whether states will go to war, because war is typically the consequence of variables that are unobservable ex ante, both to us as researchers and to the participants.[2] Thinking probabilistically continues to offer the opportunity to assess international conflict empirically. However, the realization that uncertainty is necessary theoretically to motivate war is much different from recognizing that the empirical world contains a stochastic element. Accepting uncertainty as a necessary condition of war implies that all other variables—however detailed the explanation—serve to eliminate gradations of irrelevant alternatives. We can progressively refine our ability to distinguish states that may use force from those that are likely to remain at peace, but anticipating wars from a pool of states that appear willing to fight will remain problematic. For example, we may achieve considerable success in anticipating crises, but our ability to predict which crises will become wars will probably prove little better than the naive predictions of random chance. The need for uncertainty to

1. Fearon 1995, 382.
2. I thank an anonymous reviewer for suggesting this phrasing of the argument.

**2**

"People are **not very good at understanding randomness**. There's much more chance out there than we think there is. While we are seeking for patterns and explanations as we look backward, we're not giving a fair shot to the explanation that many things are really just random events."
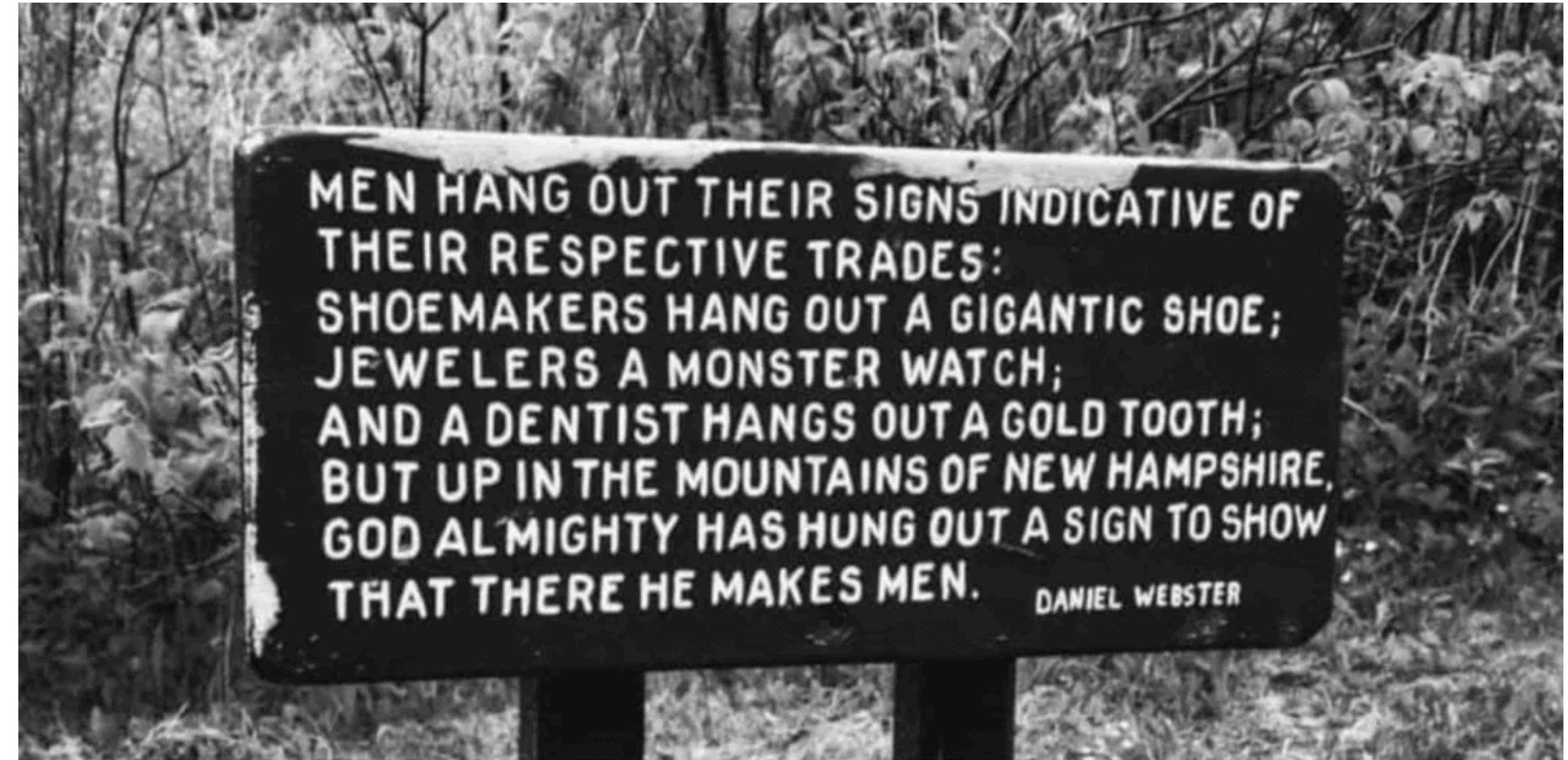
— Lisa Goldberg



Image source: https://earthsky.org/human-world/seeing-things-that-arent-there/

**Pareidolia**—Seeing things that are not there.

**Pareidolia**—Seeing things that are not there.





MEN HANG OUT THEIR SIGNS INDICATIVE OF THEIR RESPECTIVE TRADES: SHOEMAKERS HANG OUT A GIGANTIC SHOE; JEWELERS A MONSTER WATCH; AND A DENTIST HANGS OUT A GOLD TOOTH; BUT UP IN THE MOUNTAINS OF NEW HAMPSHIRE, GOD ALMIGHTY HAS HUNG OUT A SIGN TO SHOW THAT THERE HE MAKES MEN. DANIEL WEBSTER

**2**

**Conceptual clarity**—Do we know what we want to measure?

**Operational reliability**—Are the measures repeatable and consistent?

**Conceptual validity**—Does the measure accurately measure the concept we are trying to measure?

**2**

**Face validity**—On its face does a measure appear to be measuring what it says it is measuring?

**Content validity**—Does a measure capture all of the systemised concept? Is anything missing? Is anything there that should not be?

**Criterion validity**—Does a measure correlate with criterion (i.e. ground truth) variables?

**Construct validity**—Do measures behave the way you theoretically expect in the wild?

**2**

# On its face does a measure appear to be measuring what it says it is measuring?

**Socialist Constitution of The Democratic People's Republic of Korea**

조선민주주의인민공화국 사회주의헌법

Copied from Naenara website (4/29/2020 2pm)

http://naenara.com.kp/main/index/en/politics?arg_val=leader3

### PREAMBLE

The Democratic People's Republic of Korea is the socialist State of Juche where the ideas of the great leader Comrade Kim Il Sung and the great leader Comrade Kim Jong Il on State building and their exploits in it are applied.

The great leader Comrade Kim Il Sung was the founder of the Democratic People's Republic of Korea and the father of socialist Korea.

### CHAPTER V. FUNDAMENTAL RIGHTS AND DUTIES OF CITIZENS

**Article 62.** The requirements for becoming a citizen of the Democratic People's Republic of Korea are defined by the Law on Nationality. A citizen is under the protection of the Democratic People's Republic of Korea regardless of domicile.

**Article 63.** In the Democratic People's Republic of Korea the rights and duties of citizens are based on the collectivist principle: "One for all and all for one."

**Article 64.** The State shall effectively guarantee the genuine democratic rights and freedoms as well as the material and cultural well-being of all its citizens. In the Democratic People's Republic of Korea the rights and freedoms of citizens shall be amplified with the consolidation and development of the socialist system.

**Article 65.** Citizens enjoy equal rights in all spheres of State and public activity.

**2**

Does a measure capture **all** of the systemised concept? Is anything **missing**? Is anything there that **should not be**?

**POLITY™ IV PROJECT**

**Political Regime Characteristics and Transitions, 1800-2018**

**Dataset Users' Manual**

**Monty G. Marshall**
Center for Systemic Peace and
Societal-Systems Research Inc

**Ted Robert Gurr**
University of Maryland (deceased)

**Keith Jaggers**
Colorado State University

| Authority Coding | Scale Weight |
|---|---|
| *Competitiveness of Executive Recruitment (XRCOMP):* | |
| (3) Election | +2 |
| (2) Transitional | +1 |
| | |
| *Openness of Executive Recruitment (XROPEN):* | |
| only if XRCOMP is Election (3) or Transitional (2) | |
| (3) Dual/election | +1 |
| (4) Election | +1 |
| | |
| *Constraint on Chief Executive (XCONST):* | |
| (7) Executive parity or subordination | +4 |
| (6) Intermediate category | +3 |
| (5) Substantial limitations | +2 |
| (4) Intermediate category | +1 |
| | |
| *Competitiveness of Political Participation (PARCOMP):* | |
| (5) Competitive | +3 |
| (4) Transitional | +2 |
| (3) Factional | +1 |

**2**

# Does a measure correlate with criterion (i.e. **ground truth**) variables?




Labor has a significant two-party preferred lead in the polls

ALP · LNP

54.8%

45.2%

Jul 19    Nov 19    Mar 20    Jul 20    Nov 20    Mar 21    Jul 21    Nov 21    Mar 22

Polling average model by Professor Simon Jackman
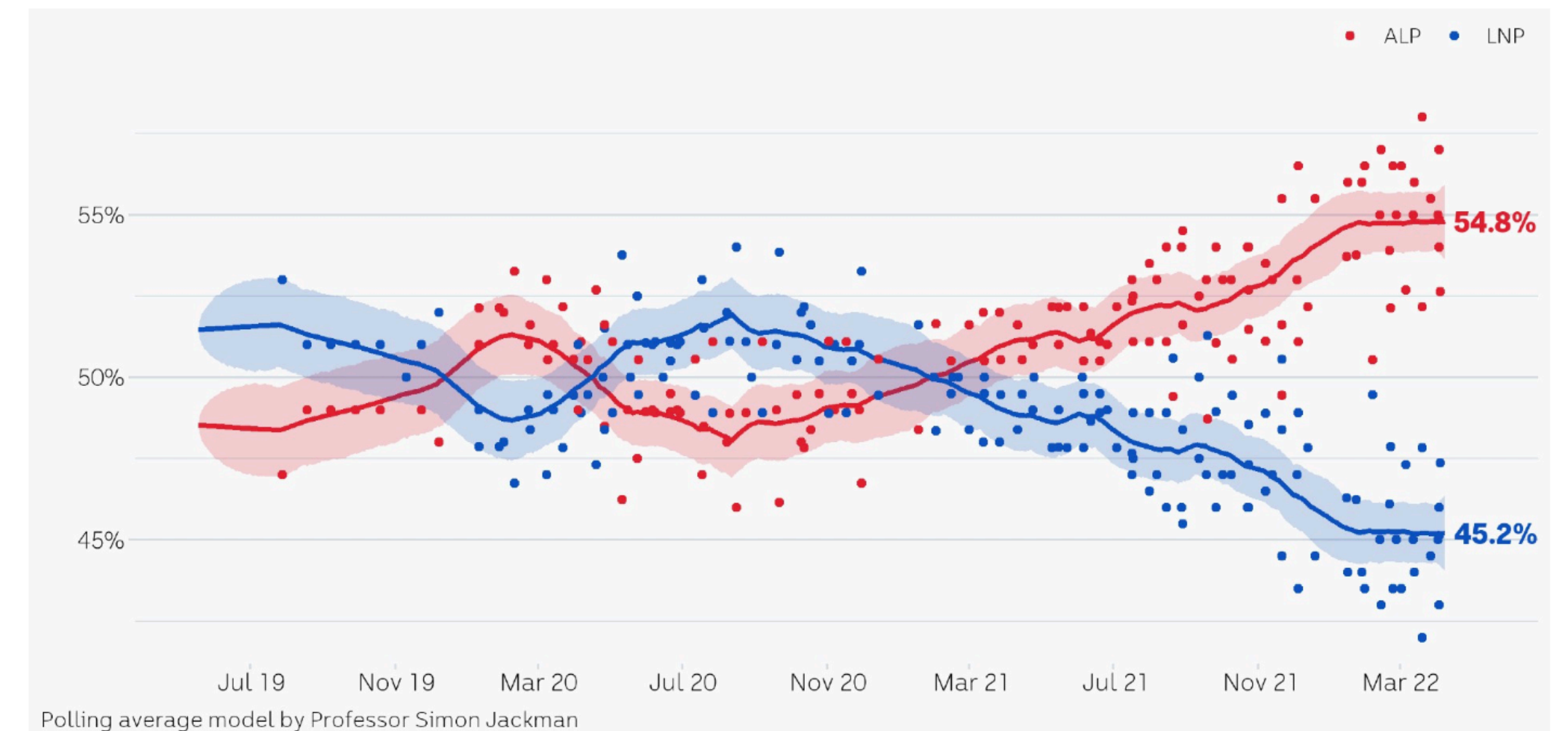
The average of the national polls since 2019

**2**

Do measures behave the way you theoretically expect **in the wild**?



Y. Tarverdi, S. Saha and N. Campbell / Economic Analysis and Policy 63 (2019) 220–233

**Fig. 2.** Governance and democracy relationship.

**2**
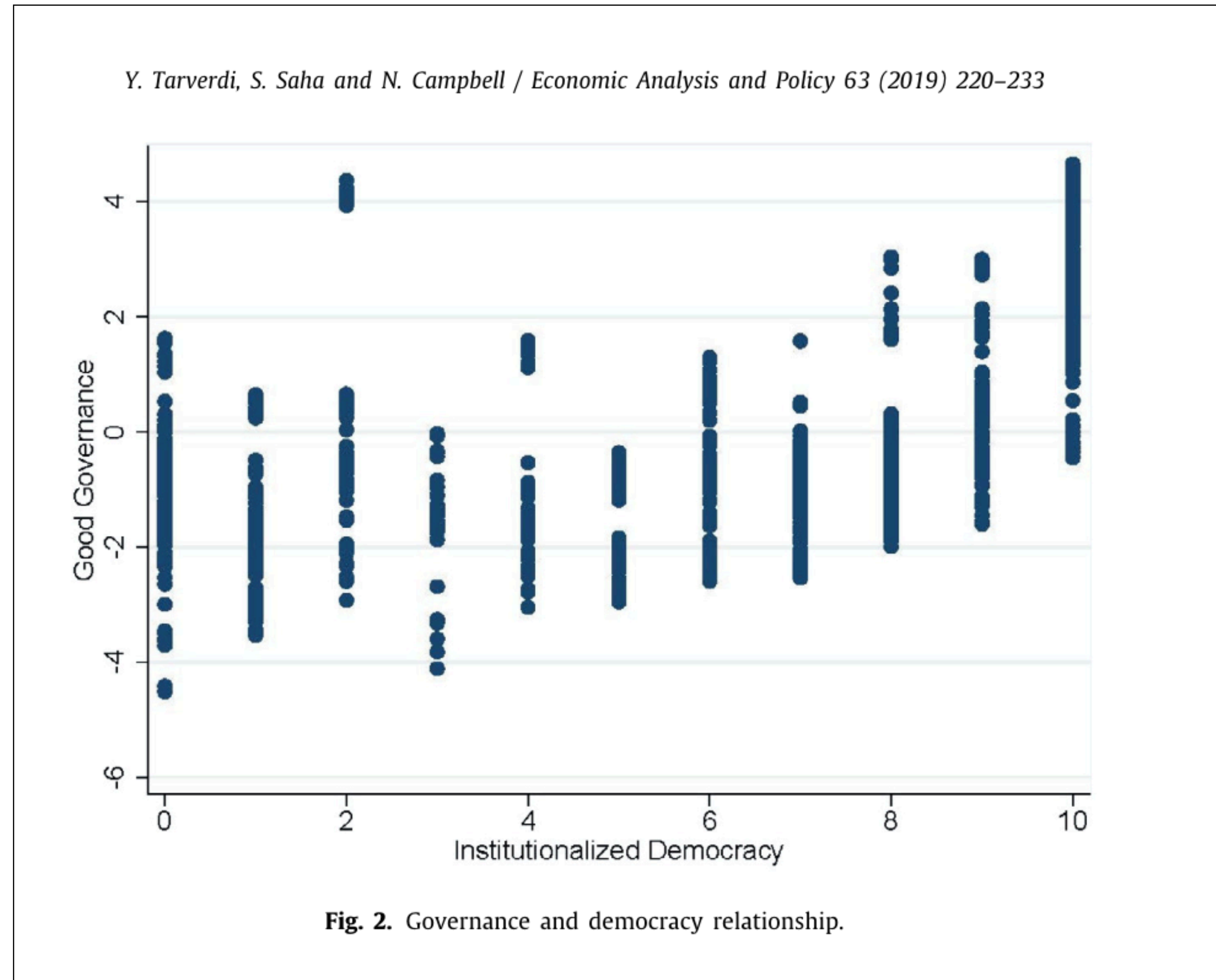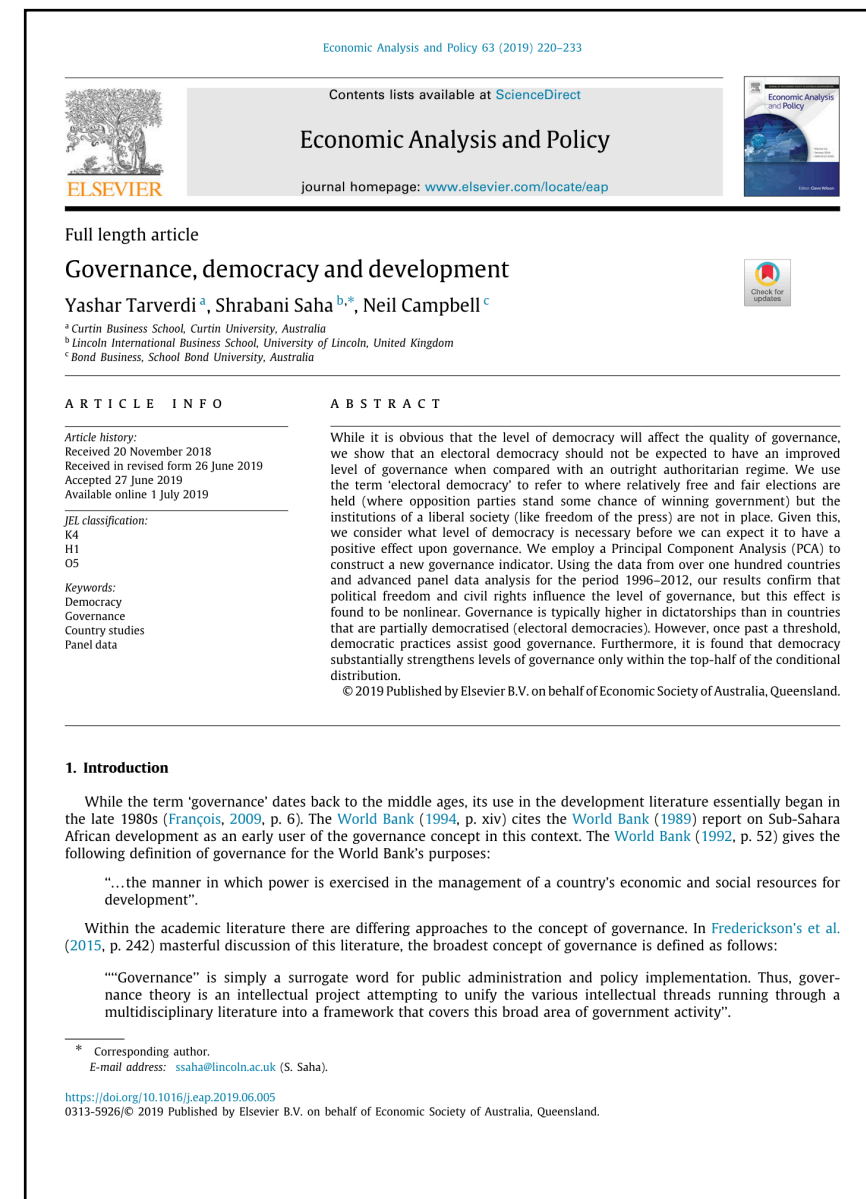
Can you think of a **valid but unreliable** measure?

A **reliable but invalid** measure?

# 3

## GDP (current US$)

GDP at purchaser's prices is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. Data are in current U.S. dollars. Dollar figures for GDP are converted from domestic currencies using single year official exchange rates. For a few countries where the official exchange rate does not reflect the rate effectively applied to actual foreign exchange transactions, an alternative conversion factor is used.

**ID**: NY.GDP.MKTP.CD

**Source**: World Bank national accounts data, and OECD National Accounts data files.

**All metadata** ⧉

2019 video source: https://youtu.be/iLom1WIqwS0

**SCIENTIFIC
AMERICAN**®

# GDP Is the Wrong Tool for Measuring What Matters

It's time to replace gross domestic product with real metrics of well-being and sustainability

By Joseph E. Stiglitz

**AUTHOR**

**Joseph E. Stiglitz** is a University Professor at Columbia University and chief economist at the Roosevelt Institute. He received the Nobel prize in economics in 2001. Stiglitz chaired President Bill Clinton's Council of Economic Advisers from 1995 to 1997 and served as the chief economist and senior vice president of the World Bank from 1997 to 2000. He chaired the Sarkozy commission (2008–2009) and an expert group (2013–2019) at the OECD for devising measures for well-being and sustainability.
*Credit: Nick Higgins*

Since World War II, most countries around the world have come to use gross domestic product, or GDP, as the core metric for prosperity. The GDP measures market output: the monetary value of all the goods and services produced in an economy during a given period, usually a year. Governments can fail if this number falls—and so, not surprisingly, governments strive to make it climb. But striving to grow GDP is not the same as ensuring the well-being of a society.

In truth, "GDP measures everything," as Senator Robert Kennedy famously said, "except that which makes life worthwhile." The number does not measure health, education, equality of opportunity, the state of the environment or many other indicators of the quality of life. It does not even measure crucial aspects of the economy such as its sustainability: whether or not it is headed for a crash. What we measure matters, though, because it guides what we do. Americans got an inkling of this causal connection during the Vietnam War, with the military's emphasis on "body counts": the weekly tabulation of the number of enemy soldiers killed. Reliance on this morbid

## Construction of the GNH Index

**The GNH Index includes nine domains**

1. Psychological wellbeing
2. Health
3. Education
4. Time use
5. Cultural diversity and resilience
6. Good governance
7. Community vitality
8. Ecological diversity and resilience
9. Living standards

Source: https://ophi.org.uk/policy/gross-national-happiness-index/
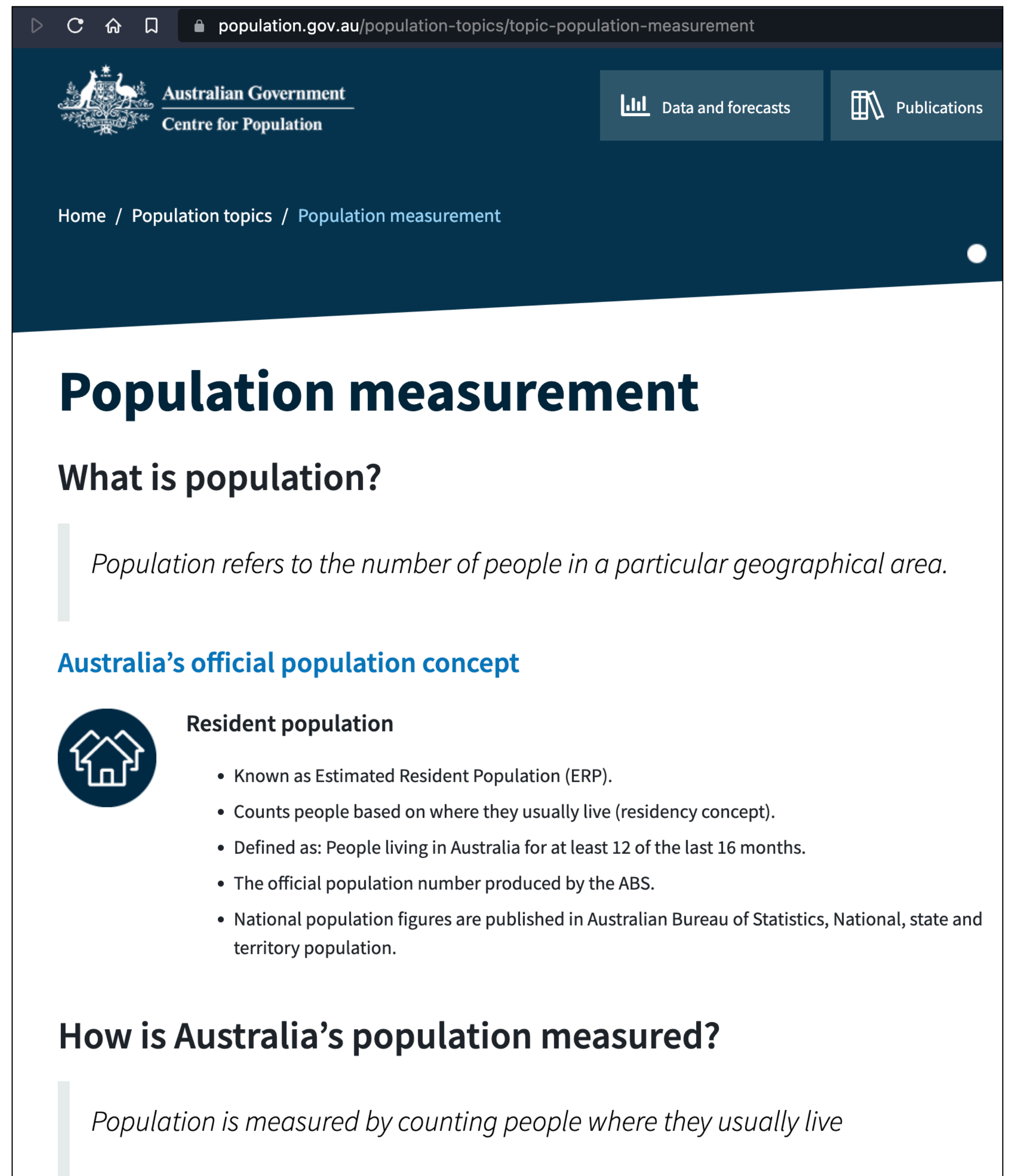
**3**

# Population, total

Total population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship. The values shown are midyear estimates.

**ID**: SP.POP.TOTL

**Source**: ( 1 ) United Nations Population Division. World Population Prospects: 2019 Revision. ( 2 ) Census reports and other statistical publications from national statistical offices, ( 3 ) Eurostat: Demographic Statistics, ( 4 ) United Nations Statistical Division. Population and Vital Statistics Reprot ( various years ), ( 5 ) U.S. Census Bureau: International Database, and ( 6 ) Secretariat of the Pacific Community: Statistics and Demography Programme.

**All metadata** ⬈

**3**



## Population measurement

### What is population?

> *Population refers to the number of people in a particular geographical area.*

### Australia's official population concept

**Resident population**

- Known as Estimated Resident Population (ERP).
- Counts people based on where they usually live (residency concept).
- Defined as: People living in Australia for at least 12 of the last 16 months.
- The official population number produced by the ABS.
- National population figures are published in Australian Bureau of Statistics, National, state and territory population.

### How is Australia's population measured?

> *Population is measured by counting people where they usually live*

**3**



Census count, actual location

**Moving people back home**
- Overseas visitors
+ residents interstate on Census night

Census count, place of usual residence

**Including missing people**
+/- people who didn't fill out the Census or were included twice
+ residents overseas

ERP as at Census night

**ERP for next quarter**
+ births
- deaths
+ arrivals
- departures

**Births and deaths** data from state and territory registries.

**Overseas arrivals and departures** data from Department of Home Affairs Travel and Immigration Processing System.

ERP as at 30 June

**Backdating from Census night to 30 June**
- births
+ deaths
- arrivals
+ departures

**Interstate arrivals and departures** data from Medicare change of address information from the Department of Health, supported by other data.

**ERP, released quarterly**

Source:https://population.gov.au/population-topics/topic-population-measurement

**Fig. (2).** Difference maps. B=EU 27+, C= GPWv3, D= GRUMP and E= Landscan.

# From Census to Grids: Comparing Gridded Population of the World with Swedish Census Records

Ola Hall[*,1], Emilie Stroh[2] and Fredy Paya[3]

[1]*Department of Human and Economic Geography, Lund University, Sweden*

[2]*Department of Occupational and Environmental Hygiene. Lund University, Sweden*

[3]*University of Waterloo, Canada*

**Abstract:** The increased availability of digital spatial data combined with improved capabilities of Geographic Information Systems (GIS) have allowed for the development of several global population distribution databases, such as the GPW, and LandScan. Making population distribution data available as a high-resolution raster database which facilitates rapid GIS analysis at the local level and for any zoning. Due to the complex nature of population as a geographical variable, several approaches have been adopted to estimate their spatial distribution, including statistical modeling, surface modeling, and cartographic methods. However, many of these methods require assumptions that over-simplify the reality or disaggregate population totals based on the heuristic or empirical parameters. Recently, critical voices were heard, questioning the quality and usability of gridded population data.

In this paper, we compare gridded population data products for parts of Sweden with high-resolution population records obtained from the Swedish National Registry through the Regional Office of Scania, Sweden. Ground-truth consists of the total population in Scania located as points at the center coordinates of their real estate (located by the Swedish Land Survey). Results indicate that there are significant differences between compared datasets.

**Keywords:** Population, GIS, census, LandScan, ground, EU+27.

## 1. INTRODUCTION

Efforts to estimate population distribution for a regular raster grid predate the computerization of geography that started in the 1980s [1]. Early examples such as the map by Adams [2] for West Africa served largely the cartographic purposes. Census offices, most notably those of Japan and Sweden, also produced national population grids for inclusion in the national atlases. Computerized population maps for individual countries were produced by the US Census Bureau using rectangular grid cells superimposed with circles for major urban areas [3]. Deichmann and Eklund hpresented a continental, gridded population database for Africa which was used to investigate interactions between population and land degradation [4]. Others, such as Martin and Bracken, developed techniques for producing local-level population grids [5].

Applications are found in various fields, from disaster management, hazards and vulnerability science, climate change science to human welfare and public health. Few studies in demography make use of this kind of data. In fact, critical voices regarding quality have been raised [6]. It seems like they are mostly used in regions with poor census records. Usually one cannot reliably assess how accurate the resulting distributions are because there is no basis for sound validation [7].

The aim of this study is to investigate the quality of four gridded population datasets, Gridded Population of the World (GPW), LandScan, Global Rural-Urban Mapping Project (GRUMP) and a recent dataset covering the European community (EU+27).
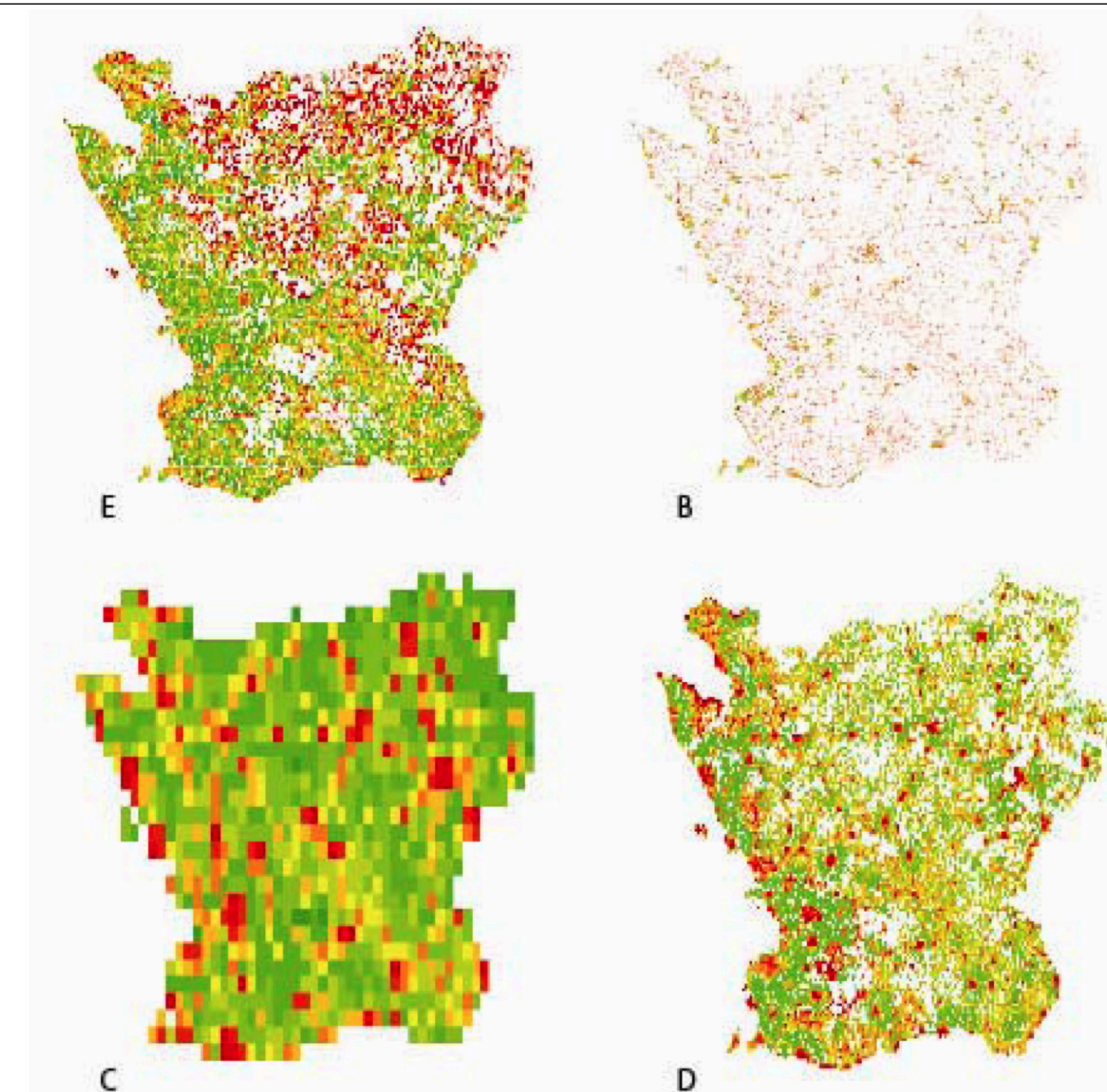
## 2. DATA AND METHODS

### 2.1. Ground Truth

Due to the structure and frequent use of personal ID numbers in Sweden the possibility of combining data at individual level from different sources is substantially better than many other countries. For example, by combining an individual's ID number with the Swedish National Land Survey's property registers the coordinates of and information on an individual's residential address can be linked to that person. The census data used in this study were obtained from the Swedish National Registry through the Regional Office of Scania, Sweden. The dataset referred to as "A" consists of the total population in Scania on the 31st of December in 2001, i.e. 1,129,059 individuals located as points at the center coordinates of their real estate (located by the Swedish Land Survey).
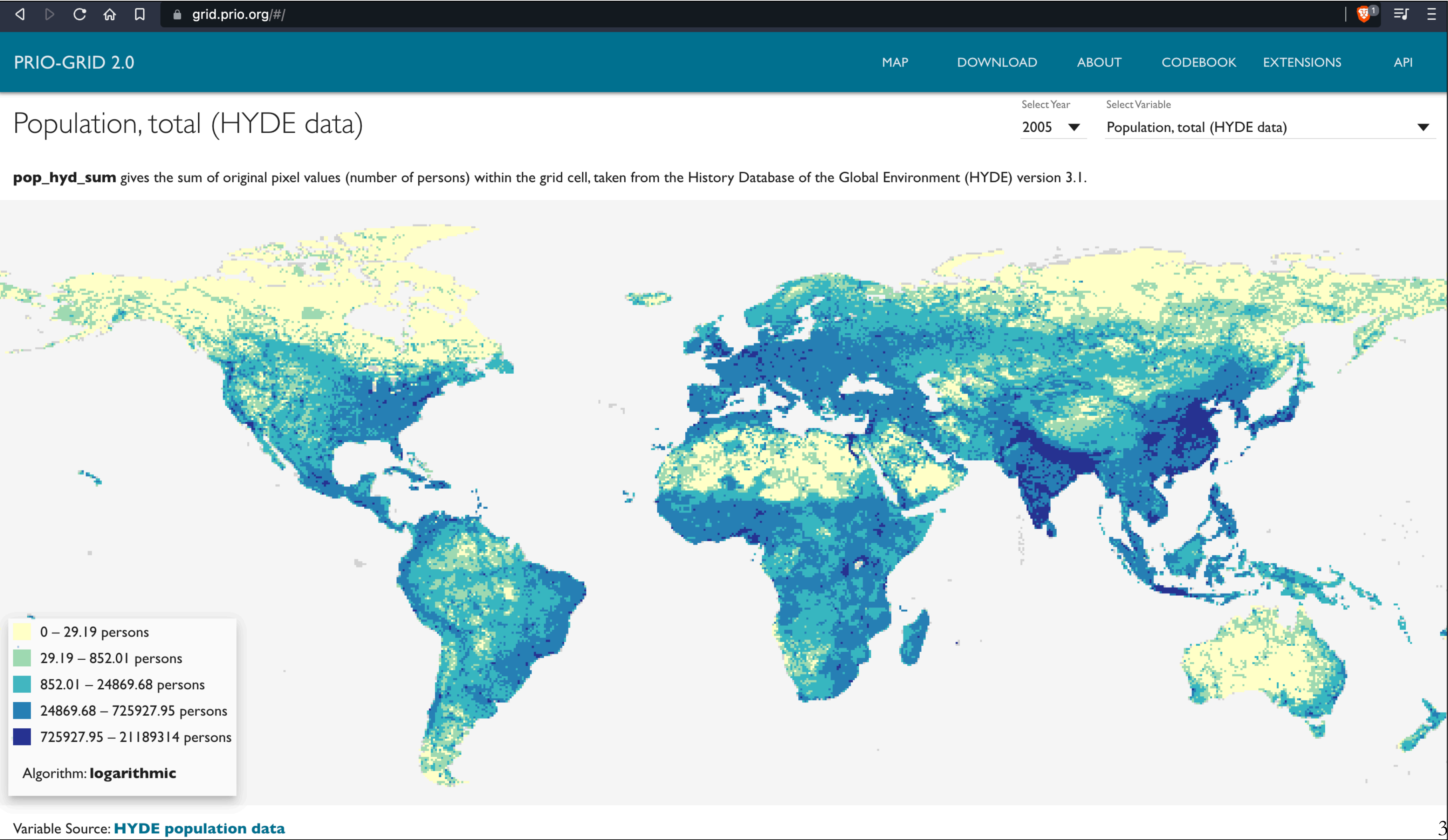
### 2.2. Population Density Grid of EU-27+

This dataset covers the 27 member states of EU and Croatia. The dataset is referred to as "dataset B" in the rest of the text. The spatial resolution is 100 x 100 m (1 ha) and in projected in Lambert-Azimuthal equal area projection. Values in the grid correspond to population density and to obtain the population count the sum of pixel values are

*Address correspondence to this author at the Department of Human and Economic Geography, Lund University, Sweden; Tel: 46 (0) 46-2220000; E-mail: ola.hall@keg.lu.se

**3**

**3**

**When** is a country a democracy?

**3**

How is democracy a
**latent/unobservable** concept?





DEMOCRACY REPORT 2022
Autocratization Changing Nature?

**3**

# V-Dem
## VARIETIES OF DEMOCRACY

## Methodology

v12 - March 2022

## 2.5 Summary

To summarize, the V-Dem conceptual scheme recognizes several levels of aggregation:

```
Core concept (1)
        Democracy Indices (5)
                Democracy Components (5)
                        Subcomponents, and related concepts (87)
                                Indicators (473)
```

# 3

Note focus on:

1. **Conceptualisation**

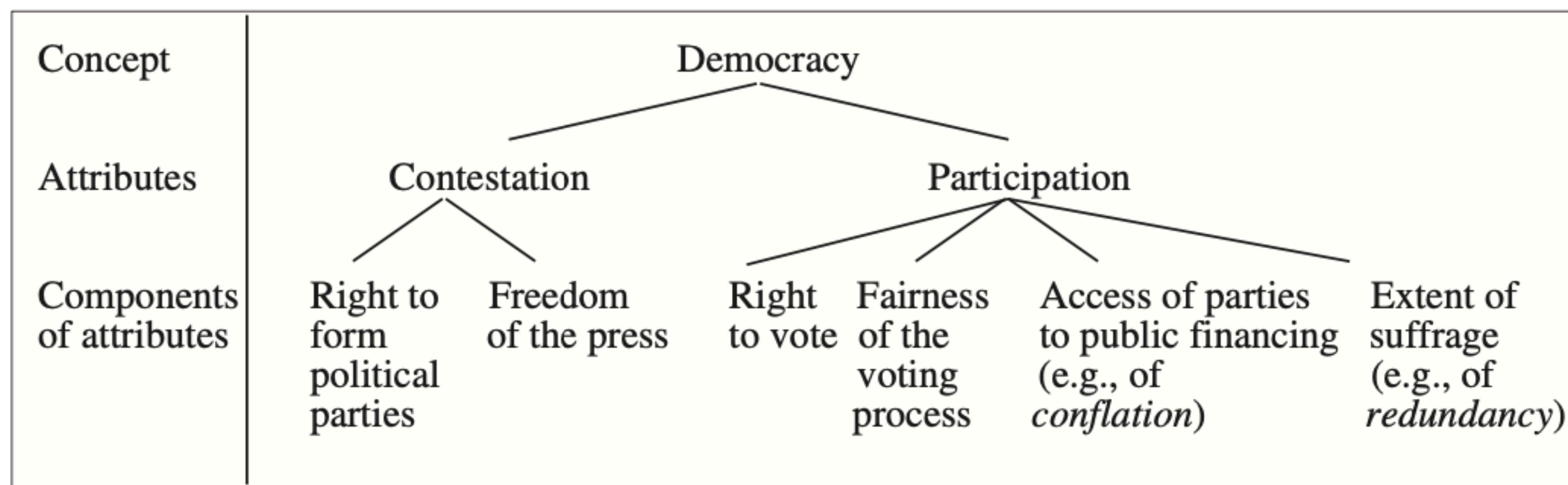2. **Measurement** (including reliability, validity, and replicability)

3. **Aggregation** (often overlooked)

Table 2

*A Framework for the Analysis of Data: Conceptualization, Measurement, and Aggregation*

| Challenge | Task | Standard of Assessment |
|---|---|---|
| Conceptualization | Identification of attributes | Concept specification: Avoid maximalist definitions (the inclusion of theoretically irrelevant attributes) or minimalist definitions (the exclusion of theoretically relevant attributes) |
| | Vertical organization of attributes by level of abstraction | Conceptual logic: Isolate the "leaves" of the concept tree and avoid the problems of redundancy and conflation |
| Measurement | Selection of indicators | Validity: Use multiple indicators and establish the cross-system equivalence of these indicators; use indicators that minimize measurement error and can be crosschecked through multiple sources |
| | | Reliability |
| | Selection of measurement level | Validity: Maximize homogeneity within measurement classes with the minimum number of necessary distinctions |
| | | Reliability |
| | Recording and publicizing of coding rules, coding process, and disaggregate data | Replicability |
| Aggregation | Selection of level of aggregation | Validity: Balance the goal of parsimony with the concern with underlying dimensionality and differentiation |
| | Selection of aggregation rule | Validity: Ensure the correspondence between the theory of the relationship between attributes and the selected rule of aggregation |
| | | Robustness of aggregate data |
| | Recording and publicizing of aggregation rules and aggregate data | Replicability |

| Concept | Democracy | | | | | |
|---|---|---|---|---|---|---|
| Attributes | Contestation | | Participation | | | |
| Components of attributes | Right to form political parties | Freedom of the press | Right to vote | Fairness of the voting process | Access of parties to public financing (e.g., of *conflation*) | Extent of suffrage (e.g., of *redundancy*) |

*Figure 1.*   **The logical structure of concepts.**
*Note*: This example has two levels of abstraction, labeled *attributes* and *components of attributes*. One could introduce a third level of abstraction, called *subcomponents of attributes*, and go even further. However, no matter how many levels of abstraction are introduced, attributes at the last level of abstraction, generically labeled as *leaves*, are used as the starting point for the task of measurement. In this example, "right to form political parties" is a leaf.

# Natural Resources, Conflict, and Conflict Resolution

## UNCOVERING THE MECHANISMS

**MACARTAN HUMPHREYS**

*Department of Political Science*
*Columbia University*

The interpretation of the resource-conflict link that has become most publicized—the rebel greed hypothesis—depends on just one of many plausible mechanisms that could underlie a relationship between resource dependence and violence. The author catalogues a large range of rival possible mechanisms, highlights a set of techniques that may be used to identify these mechanisms, and begins to employ these techniques to distinguish between rival accounts of the resource-conflict linkages. The author uses finer natural resource data than has been used in the past, gathering and presenting new data on oil and diamonds production and on oil stocks. The author finds evidence that (1) conflict onset is more responsive to the impacts of past natural resource production than to the potential for future production, supporting a weak states mechanism rather than a rebel greed mechanism; (2) the impact of natural resources on conflict cannot easily be attributed entirely to the weak states mechanism, and in particular, the impact of natural resources is independent of state strength; (3) the link between primary commodities and conflict is driven in part by agricultural dependence rather than by natural resources more narrowly defined, a finding consistent with a "sparse networks" mechanism; (4) natural resources are associated with shorter wars, and natural resource wars are more likely to end with military victory for one side than other wars. This is consistent with evidence that external actors have incentives to work to bring wars to a close when natural resource supplies are threatened. The author finds no evidence that resources are associated with particular difficulties in negotiating ends to conflicts, contrary to arguments that loot-seeking rebels aim to prolong wars.

*Keywords:* civil war; resource-conflict link; resource dependence

## 1. INTRODUCTION

In early April 1975, President Tombalbaye of Chad appealed on national radio for popular vigilance, warning that members of the army were plotting a coup against him. He explained that if anyone wanted to know *why* a coup was being plotted, the answer lay with the oil in the Doba fields in the south. This appeal turned out to be Tombalbaye's last public address. On April 13, he was killed during Chad's first suc-

**3**

1. Greedy rebels

2. Greedy outsiders

3. Grievances

4. Feasibility

5. Weak states

6. Sparse networks

1. Diamond production

2. State instability

3. State instability x autocracy

4. Oil reserves (per capita)

5. Oil production (per capita)

6. share of agriculture (% GDP)
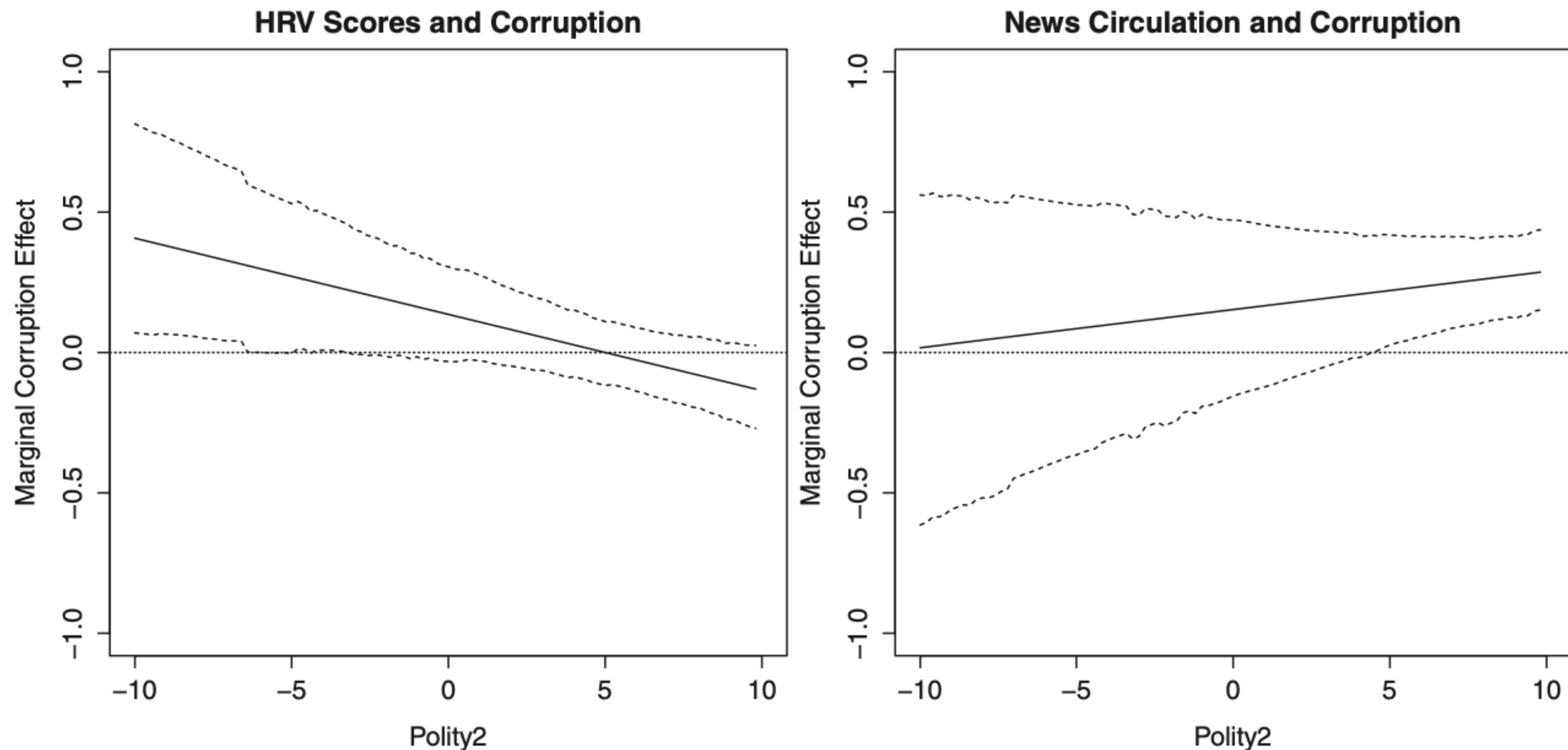
**Fig. 13** Marginal effects of a standard deviation change in transparency on corruption. Plots of the marginal effect of transparency on corruption for varying levels of democracy. Polity2 (democracy) scores are plotted on the *x*-axis; the marginal effect of a 1 SD change in each transparency measure is plotted on the *y*-axis. Point estimates are represented by the solid line; 95% highest posterior density intervals are represented by the dashed lines.

# Measuring Transparency

**James R. Hollyer**
*Department of Political Science, University of Minnesota, 1414 Social Sciences, 267 19th Ave. S, Minneapolis, MN 55455*
*e-mail: jhollyer@umn.edu (corresponding author)*

**B. Peter Rosendorff**
*Wilf Family Department of Politics, New York University*
*e-mail: peter.rosendorff@nyu.edu*

**James Raymond Vreeland**
*Mortara Center for International Studies, Georgetown University*
*e-mail: jrv24@georgetown.edu*

Edited by Jonathan Katz

Transparency is often viewed as crucial to government accountability, but its measurement remains elusive. This concept encompasses many dimensions, which have distinct effects. In this article, we focus on a specific dimension of transparency: governments' collection and dissemination of aggregate data. We construct a measure of this aspect of transparency, using an item response model that treats transparency as a latent predictor of the reporting of data to the World Bank's World Development Indicators. The resultant index covers 125 countries from 1980 to 2010. Unlike some alternatives (e.g., Freedom House), our measure—the HRV index—is based on objective criteria rather than subjective expert judgments. Unlike newspaper circulation numbers, HRV reflects the dissemination of credible content—in that it has survived the World Bank's quality control assessment. In a validation exercise, we find that our measure outperforms newspaper circulation as a predictor of Law and Order and Bureaucratic Quality as measured by the ICRG, particularly in autocracies. It performs as well as newspaper circulation in predicting corruption. These findings suggest that data dissemination is a distinct, and politically relevant, form of transparency.

Transparency, as broadly defined, relates to the full flow of information within a polity. A lengthy literature stresses the importance of the provision of information to the populace to ensure the accountability of government, and thus emphasizes the role of transparency as a determinant of government performance (see, for instance, Adserà, Boix, and Payne 2003; Besley and Burgess 2002).

This broad notion of transparency may encompass many different forms of information transmission. Common proxies for transparency measure aspects of the media market, the presence/absence of freedom of information laws (FOILs), or even political institutions such as democracy. A smaller portion of the literature considers an alternative dimension of transparency: the government's collection and dissemination of aggregate economic data.

In this article, we introduce a novel measure of this last dimension of transparency. To construct this measure, we rely on a Bayesian item response (IRT) model, which treats transparency (along

**3**

How can we link solid **causal theories** to **real-world evidence**?

How can we be sure this evidence has **measurement validity**?

**3**

## Important terms

Chapter 5