



UNIVERSITY OF  
BIRMINGHAM

COLLEGE OF  
SOCIAL SCIENCES

ESRC  
FESTIVAL  
OF SOCIAL  
SCIENCE

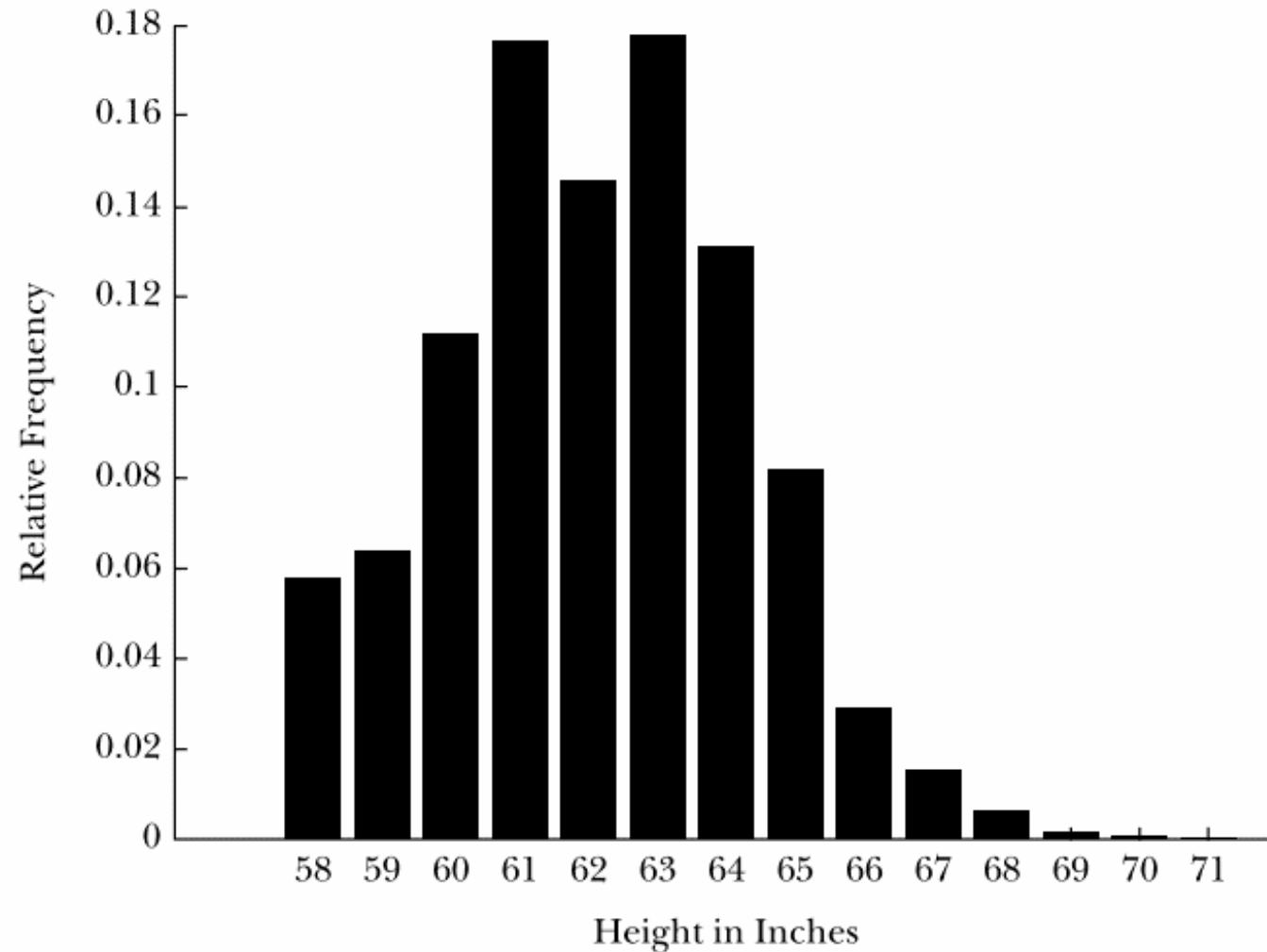
ESRC FESTIVAL OF SOCIAL SCIENCE



## Good Data Gone Bad: playing fast and loose with data

Colin Rowat

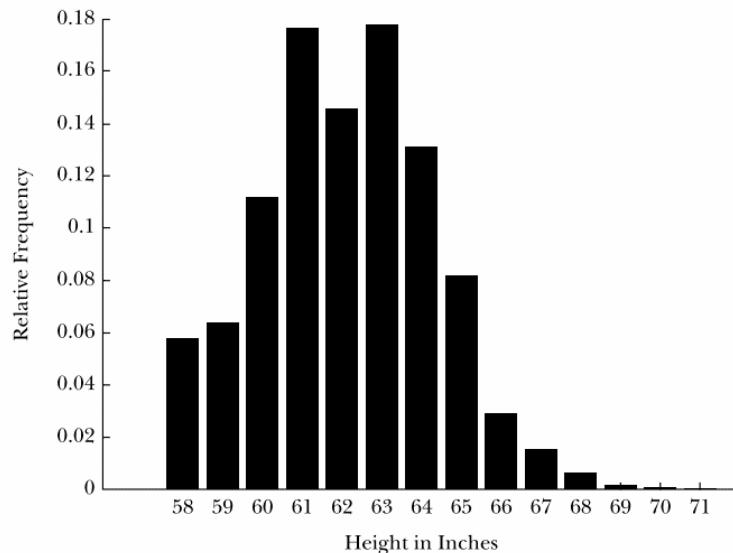
## Doubts About the Doubs: Bertillon's Histogram of the Heights of 9,002 Conscripts from Doubs



# How do we explain this bimodal distribution?

- the bimodal distribution survives slicing of the data: e.g. recruits from 1851 – 1855 or from 1856 – 1860
- Aha! Doubs is inhabited by two different ‘races’:
  - **Celts:** “tall, healthy and ‘true to their word’”
  - **Burgundians:** “feckless ... shorter persons neither ‘so robust, nor temperate, nor obliging’”

Doubts About the Doubs: Bertillon’s Histogram of the Heights of 9,002 Conscripts from Doubs



- 26 years later, an alternative explanation...

“Livi noted that the ‘dip’ that appears in Bertillon’s histogram at the bin containing the relative frequency of conscripts with heights between 5’1” and 5’2” contains only *two* centimeter classes, while the bins to the left and right, from 5’0” to 5’1” and from 5’2” to 5’3”, contained *three* such centimeter classes.”

# Growth in a time of debt (Reinhart & Rogoff, AER 2010)

- two centuries of data: 1790 – 2009
- average growth rates against debt/GDP ratios

“When gross external debt reaches 60% of GDP, annual growth declines by about 2%; for levels of external debt in excess of 90% of GDP, growth rates are roughly cut in half”

- used to support austerity policies in US, EU, UK
- Herndon, Ash and Pollin (2013) find Excel errors
  - missing data on several countries
  - missing some years
  - they find a smaller effect
- no insight into direction of causality

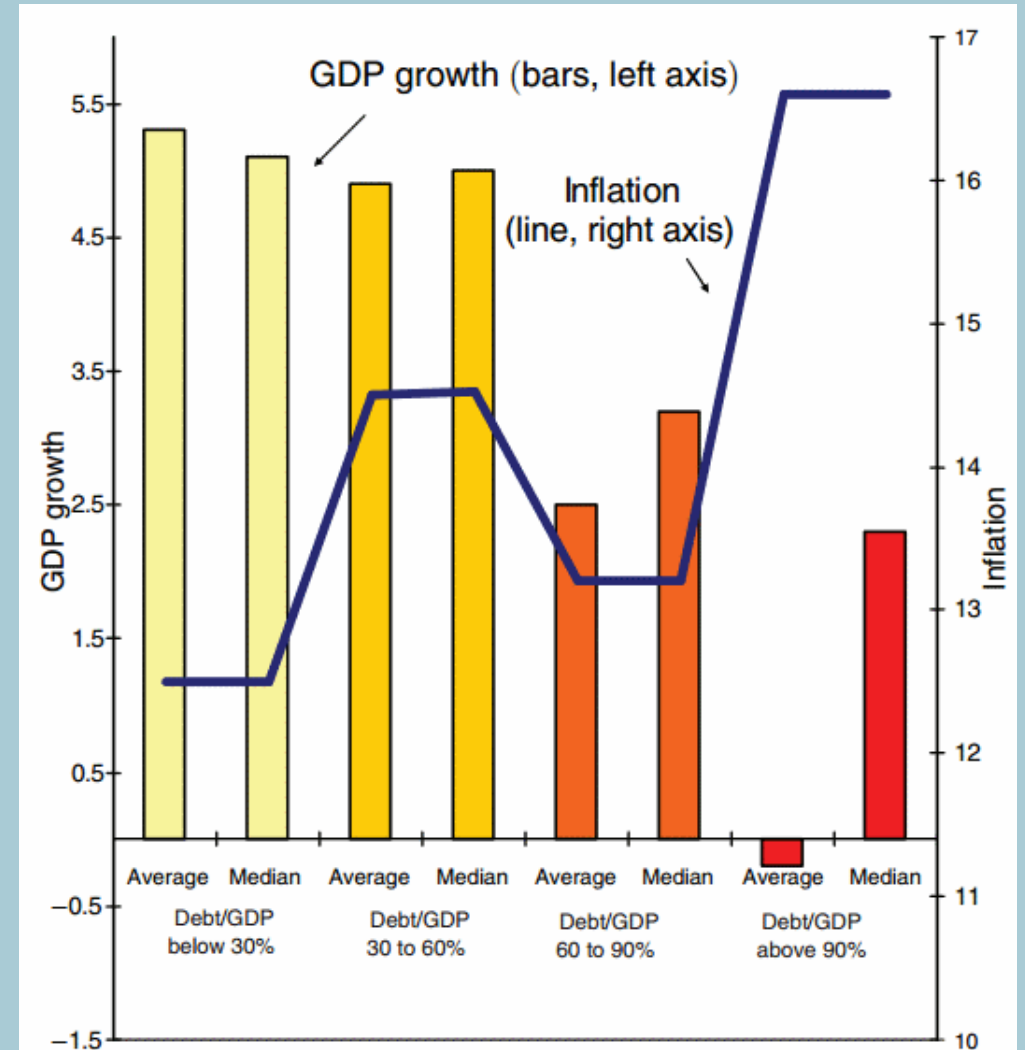
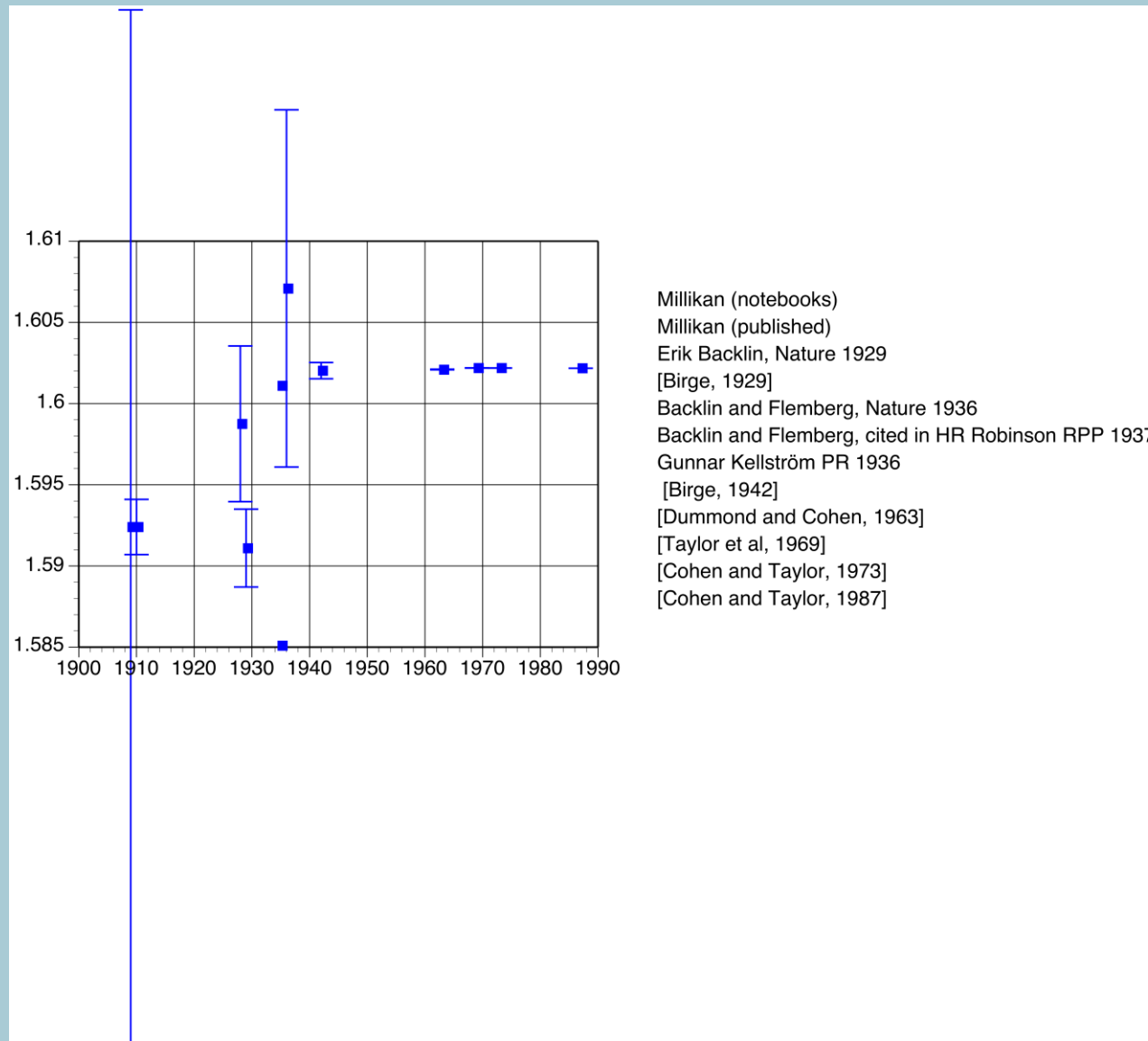


FIGURE 3. EXTERNAL DEBT, GROWTH, AND INFLATION: SELECTED EMERGING MARKETS, 1970-2009

# Changes in electron charge measurements



# Explaining the slow convergence of electron charge results

- hint: Millikan had won the Nobel Prize for Physics in 1923 for this work

“Why didn't they discover the new number was higher right away? ... it's apparent that people did things like this: When they got a number that was too high above Millikan's, they thought something must be wrong – and they would look for and find a reason why something might be wrong. When they got a number close to Millikan's value they didn't look so hard. And so they eliminated the numbers that were too far off, and did other things like that.” (Feynman, 1974)

- some branches of physics ‘blind’ their data
  - prevent theoretical priors influencing interpretation

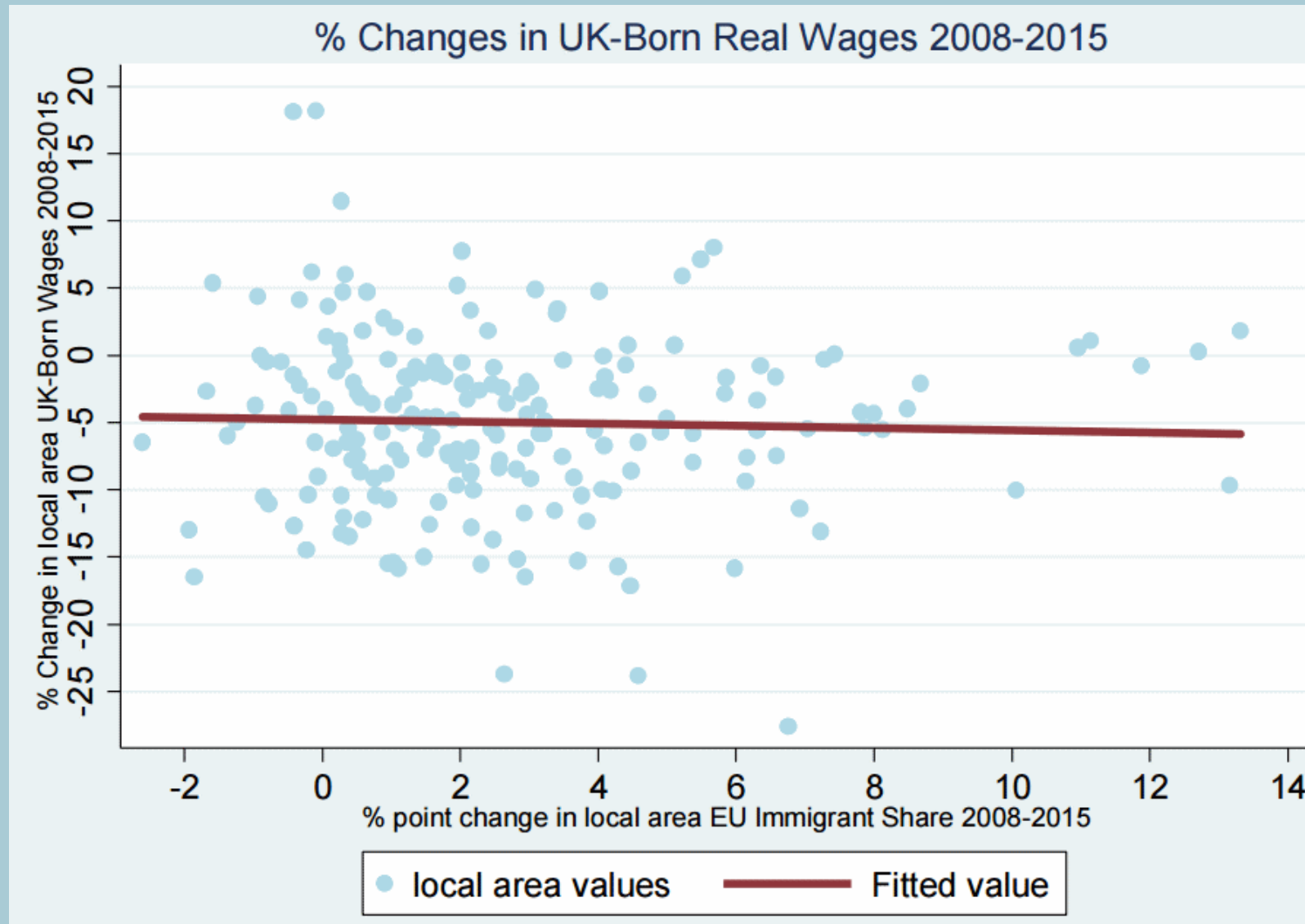


# Does 'big data' eliminate the need for theory?

“[Big data is] absolutely *not* an end to theory. In fact, the need for theory is in some ways magnified by having large amounts of data. When you have a small amount of data, you can just look at the data and build your intuition from it. When you have very large amounts of data, just taking an average can cost thousands of dollars of computer time. So you'd better have an idea of what you're doing and why before you go out to take those averages. The importance of theory to create conceptual frameworks to know what to look for has never been larger, I think.” (Athey, 2013)



# EU migration has no effect on UK-born wages

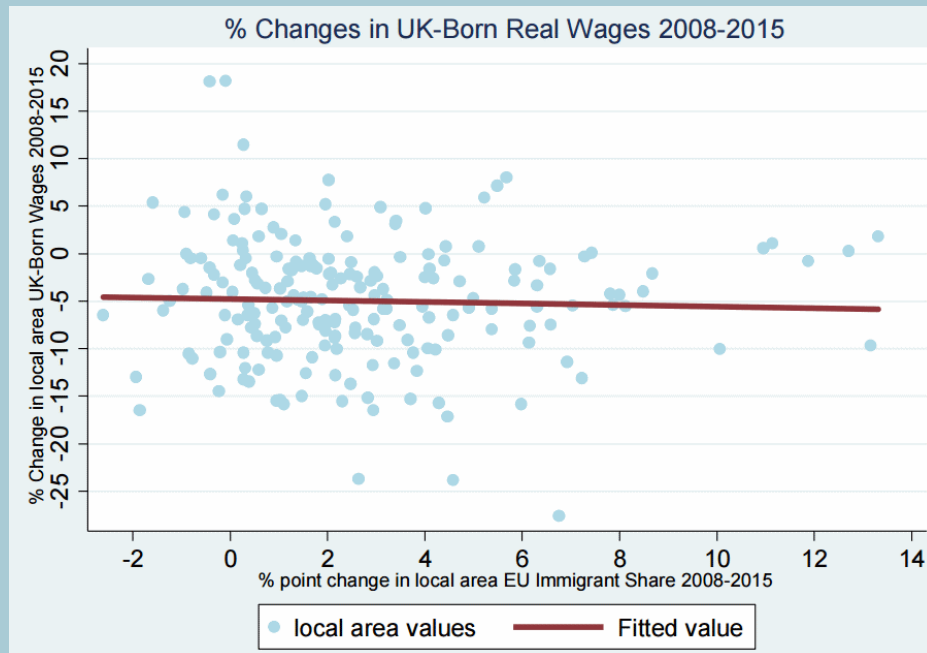




# Does lack of correlation imply lack of causation?

- suppose that real wages grow in regions of economic prosperity
- this, in turn, attracts migrants
- if EU migration didn't depress wages, would we not expect to see a positive correlation?

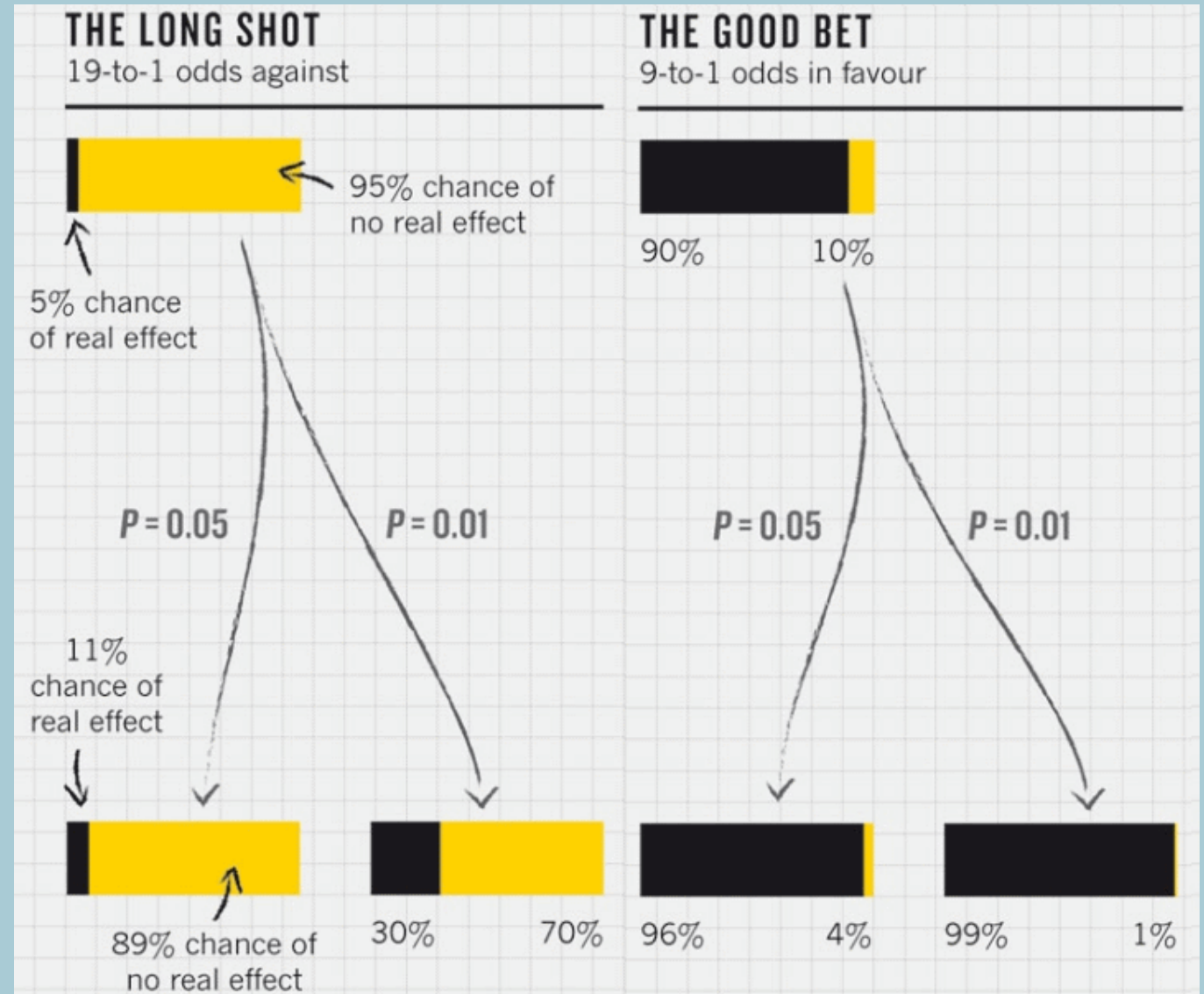
(Stephen Wright, 2016)



# But my p-values are highly significant...

“a study of nearly 2,000 people seemed to show that political moderates saw shades of grey more accurately than did ... extremists. ... The  $P$  value ... was 0.01 – usually interpreted as ‘very significant’. ... Sensitive to controversies over reproducibility, Motyl and ... Nosek, decided to replicate the study. With extra data, the  $P$  value came out as 0.59 — not even close to the conventional level of significance, 0.05.” Nuzzo (2014)

- 2/3 of economics’ results replicable (Camerer et al., 2016)?

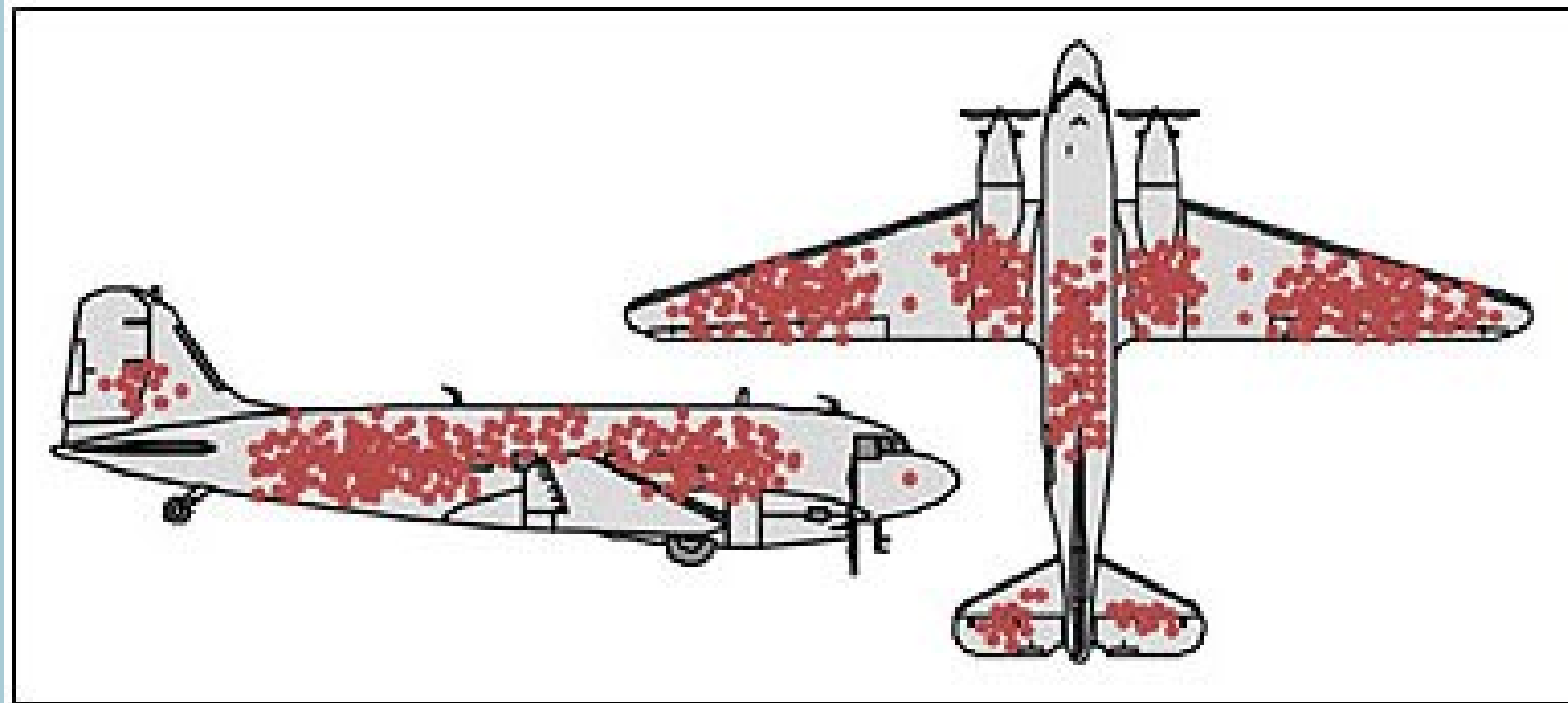


# These issues are found across disciplines

- sanctions's cost potentially huge (e.g. Iraq's child mortality rate): do they work?
- Hufbauer, Schott & Elliott (1990) produced most widely cited figure
  - 16-point score for each major sanctions episode
  - 9 or higher is a 'success'
  - conclusion: of 116 major cases, sanctions work in 34%
- But...
  - omitted variables bias (e.g. support for the sanctions)
  - correlation v. causation (why are the parties in conflict in the first case?)
  - imposing sanctions means that something's gone wrong (Eaton & Engers, 1999)
    - Drezner (2003): threat of sanctions more effective than their use



# Clear thinking leads to theoretical developments



Credit: Cameron Moll

- where would you increase the armour on WWII bombers?
- US Center for Naval Analysis: put it where the planes suffer most damage
- Wald & selection bias: the holes show you where a plane can get hit and still survive