Toulouse School of Economics

Applied Econometrics for Development

TSE M2 – Semester 1

September 2019

Ana Gazmuri and Paul Seabright

Week 1 (September 23rd): Introduction

Economics for the Common Good

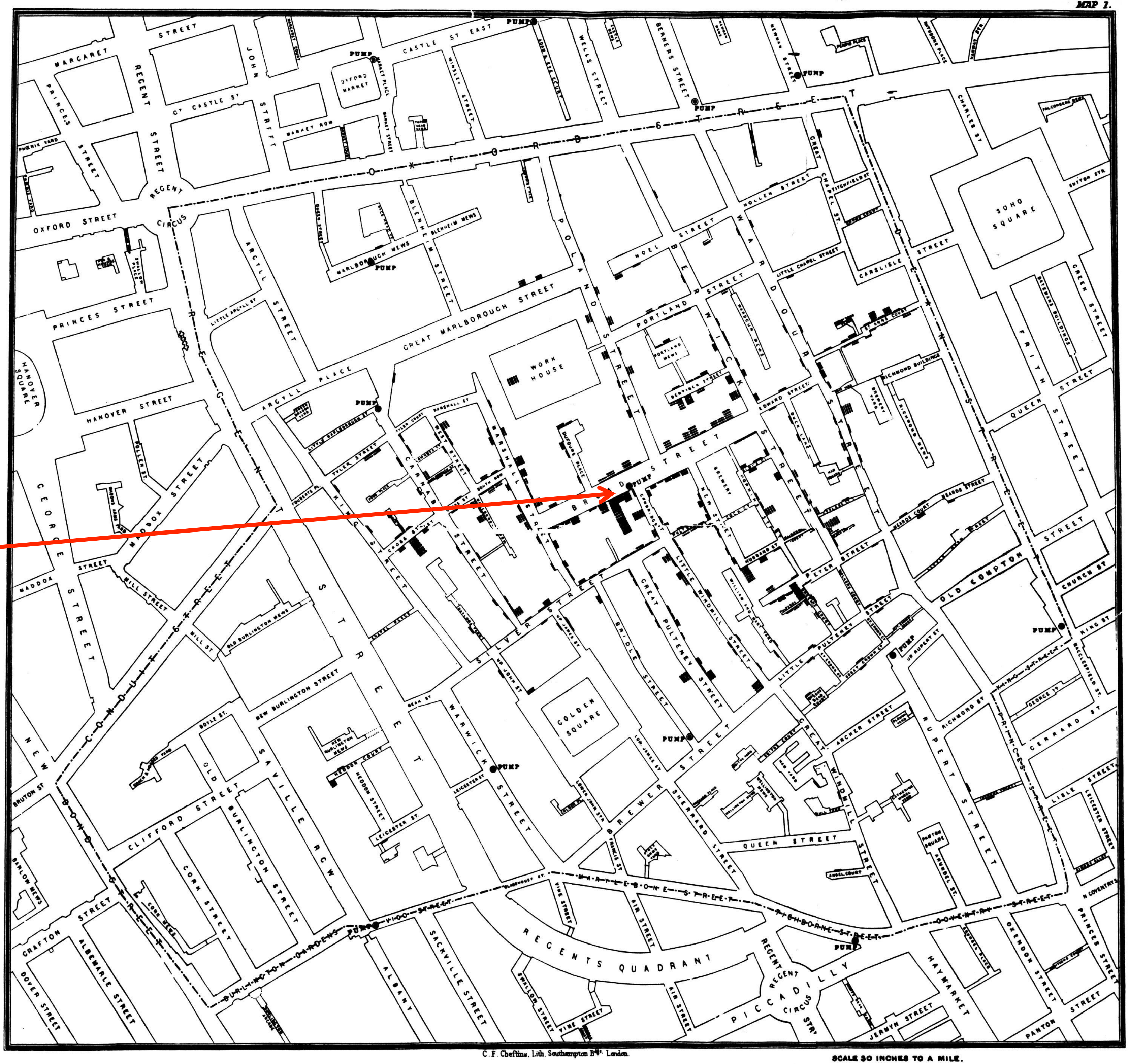
An early example of statistics in public policy: physician John Snow and the London cholera outbreak of 1854

- On proceeding to the spot, I found that nearly all the deaths had taken place within a short distance of the [Broad Street] pump. There were only ten deaths in houses situated decidedly nearer to another street-pump. In five of these cases the families of the deceased persons informed me that they always sent to the pump in Broad Street, as they preferred the water to that of the pumps which were nearer. In three other cases, the deceased were children who went to school near the pump in Broad Street... With regard to the deaths occurring in the locality belonging to the pump, there were 61 instances in which I was informed that the deceased persons used to drink the pump water from Broad Street, either constantly or occasionally...

- The result of the inquiry, then, is, that there has been no particular outbreak or prevalence of cholera in this part of London except among the persons who were in the habit of drinking the water of the above-mentioned pump well.

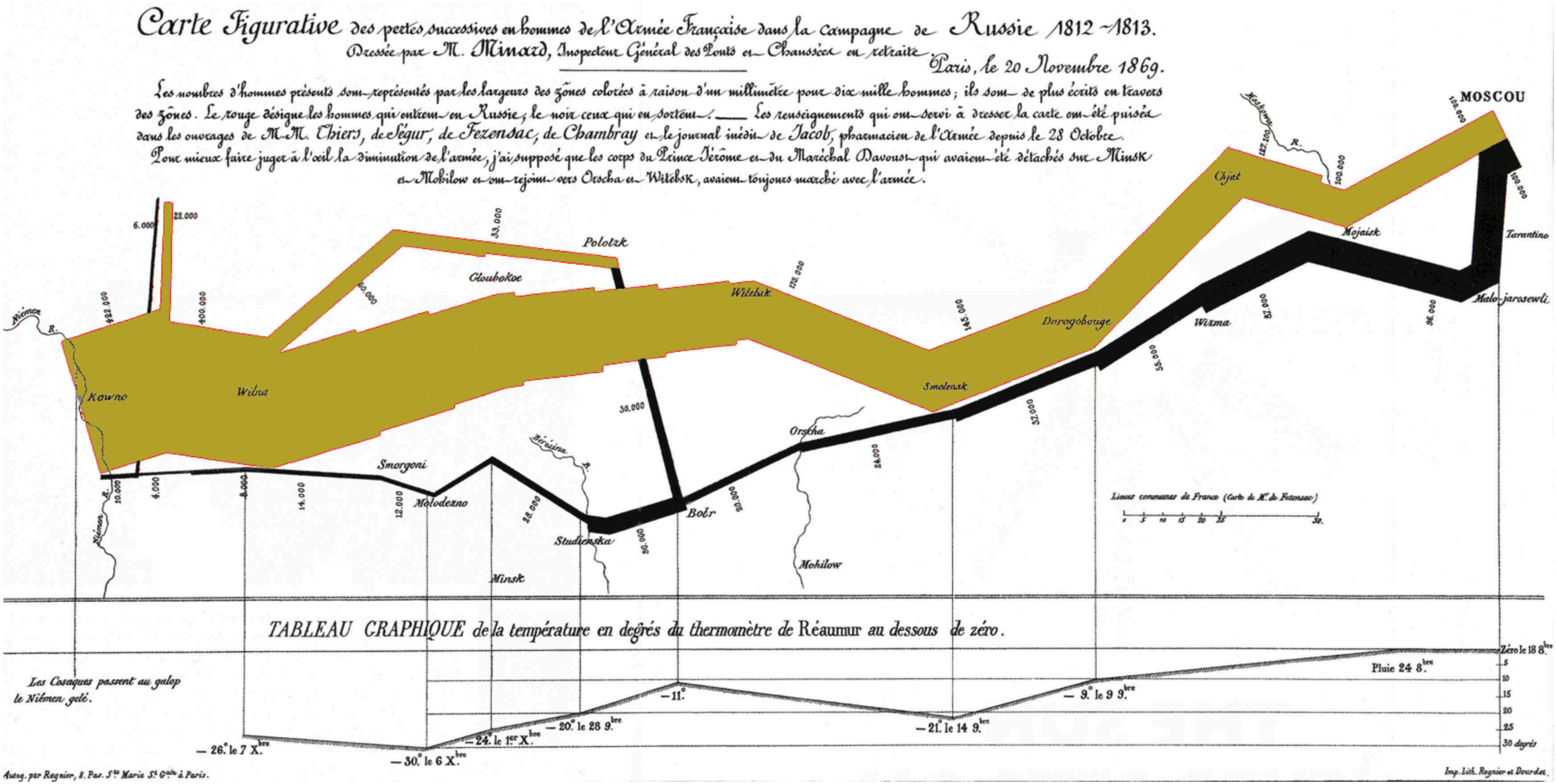
- I had an interview with the Board of Guardians of St James's parish, on the evening of the 7th inst [Sept 7], and represented the above circumstances to them. In consequence of what I said, the handle of the pump was removed on the following day.

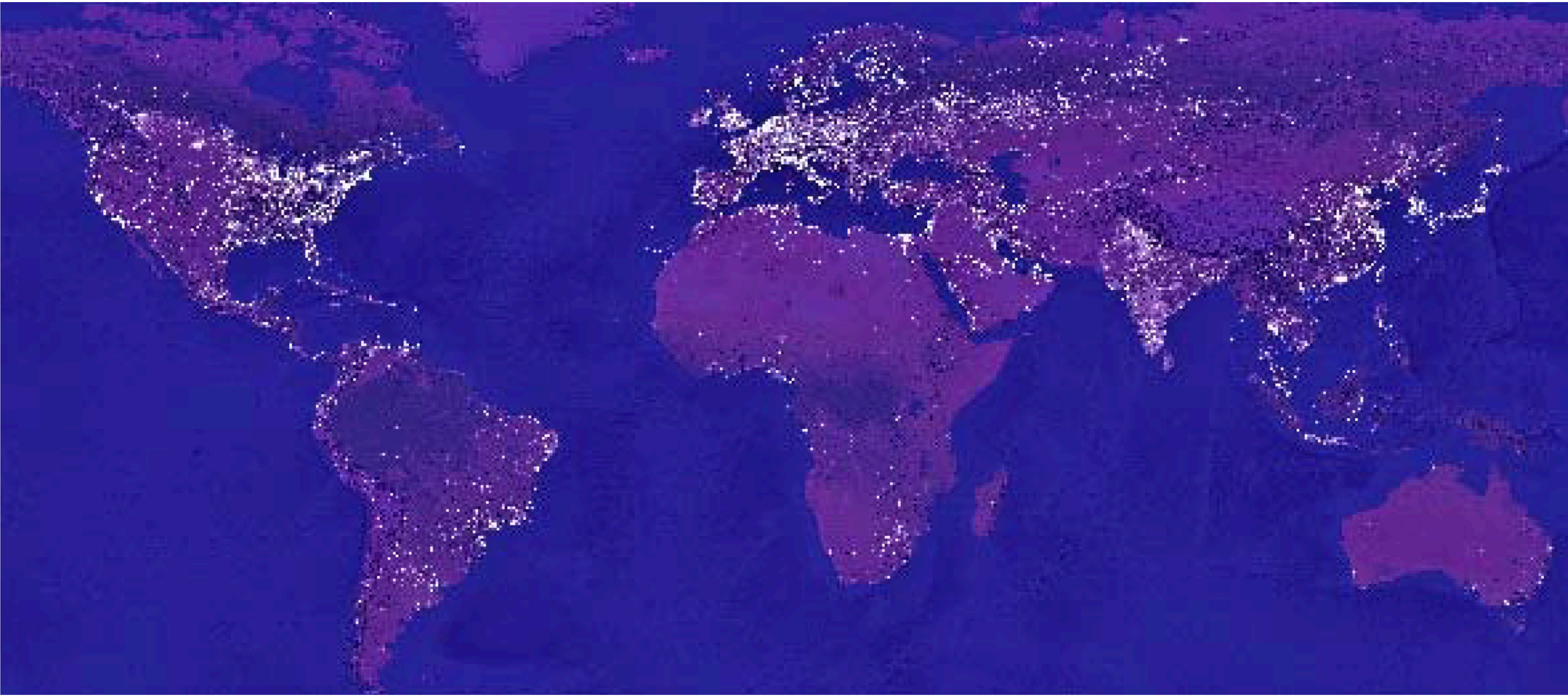
The Broad Street pump



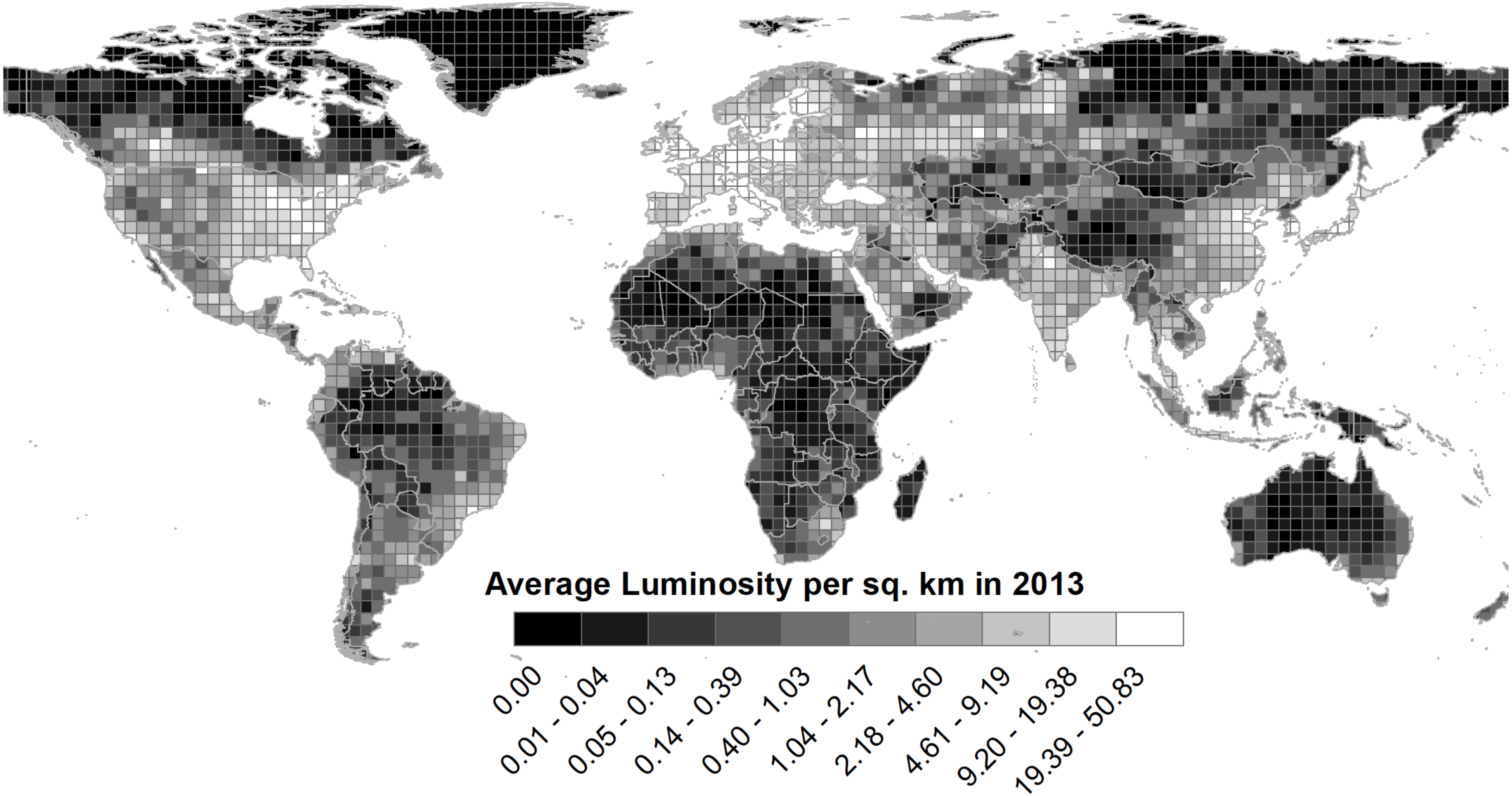
Not all statistical presentations have to be in the form of graphs and regression tables…

- There are many other well known examples:
- Charles Joseph Minard (1781-1870) drew one of the most effective statistical graphs ever made
- Nowadays spatial representations are possible using geo-located data


Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812–1813.
Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite.
Paris, le 20 Novembre 1869.
Les nombres d'hommes présents sont représentés par les largeurs des zônes colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zônes. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M.M. Thiers, de Ségur, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.
Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davoust qui avaient été détachés sur Minsk et Mobilow et ont rejoint vers Orscha et Witebsk, avaient toujours marché avec l'armée.
MOSCOU
Tarantino
Malo-jarosewli
Mojaisk
Chjat
Wixma
Dorogobouge
Smolensk
Orscha
Mohilow
Bobr
Studienska
Minsk
Molodezno
Smorgoni
Witebsk
Polotzk
Gloubokoe
Wilna
Kowno
Niémen R.
Bérézina R.
422.000
400.000
22.000
6.000
60.000
33.000
175.000
145.000
127.100
100.000
100.000
96.000
87.000
55.000
37.000
24.000
20.000
30.000
50.000
28.000
12.000
14.000
8.000
4.000
10.000
Lieues communes de France (Carte de M.r de Fezensac)
TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.
Les Cosaques passent au galop le Niémen gelé.
Zéro le 18 8.bre
Pluie 24 8.bre
— 9.° le 9 9.bre
— 11.°
— 21.° le 14 9.bre
— 20.° le 28 9.bre
— 24.° le 1.er X.bre
— 30.° le 6 X.bre
— 26.° le 7 X.bre
30 degrés
Autog. par Regnier, 8. Pas. S.te Marie S.t G.ain à Paris.
Imp. Lith. Regnier et Dourdet.



Source: Michalopoulos & Papaioannou 2017, Figure 1



Source: Michalopoulos & Papaioannou 2017, Figure 7

Another example: the prevention of scurvy

- Scurvy is a disease caused by deficiency of Vitamin C, leading to bleeding gums and skin and general lethargy, and eventually death. Known since Hippocrates (5th century BCE), it has been subject of folk remedies involving citrus fruits and other ingredients (eg bark).

- Scurvy is estimated to have killed 2 million sailors between 1500 and 1800 (Drymon 2008), more than were killed by enemy action. In the Seven Years War, 133,000 out of 189,000 Royal Navy conscripts died of disease, mainly scurvy

- In 1753 James Lind published *A Treatise on the Scurvy* detailing the first ever clinical trial, using citrus fruits and contrasting the effects with those of other acid foods

Outline of Presentation

- The objective: causal investigation of economic processes
- The fundamental difficulty: the unobservability of outcomes under alternative interventions
- Underlying motivation – the p-values controversy
- A sketch of three types of approach
  - Experiments
  - Instrumental variables
  - Panel data

Define *potential outcomes*

Let i be an individual

Y_i be an outcome of interest

W_i be a causal intervention to be investigated

Then $Y_i = Y_i(W_i) = Y_i(0)(1 - W_i) + Y_i(1)W_i$

$$= \begin{cases} Y_i(0) & \text{if } W_i = 0 \\ Y_i(1) & \text{if } W_i = 1 \end{cases}$$

Defining the causal (or treatment) effect

Treatment effect of W for i is $Y_i(1) - Y_i(0)$

This is a relation between an observable and an unobservable.

Compare the regression equation

$$(1) \qquad Y_i = \alpha + \beta . W_i + \varepsilon_i$$

The parameter estimate $\hat{\beta}$ captures a relation between observables only

What to do?

- In some way or other, observed outcomes have to be used to estimate the characteristics of potential but unobserved outcomes.

- The main difficulty: Potential outcomes may vary between individuals

$$\underbrace{Y_i(1) - Y_j(0)}_{\text{Measured effect}} \neq \underbrace{Y_i(1) - Y_i(0)}_{\text{True effect}} \text{ if } Y_j(0) \neq Y_i(0)$$

A classic illustration at a macro level

- A country undergoes a political change (democratization, a Leninist revolution)
- How to measure what would have happened in its absence?
- This means we have difficulty even defining the nature of the treatment
- Example: North and South Korea: Cowen and Tabarrok (2012) say « broadly speaking, South Korea chose capitalism, North Korea chose communism, and the results are so clear you could see them from outer space »

The Importance of Institutions

What's wrong with the use of this comparison?

- That the two countries had a divergence in institutions is undoubted
- But the « treatment » was not just the « standard » capitalism versus « standard » communism.
- In statistical terms both countries are outliers in their respective groups
- How to obtain the correct treatment effect of central planning?

Underlying motivation: the p-values controversy

- In recent years there has been growing controversy over the excessive importance attached to p-values, especially the 5% significant threshold.

- The editors of the American Economic Review have announced that they no longer wish authors to include asterisks for significance levels in their tables

- Criticisms of the 5% significance threshold:
  - It encourages p-hacking.
  - It fails to distinguish between statistical and economic significance.
  - It treats as a discrete phenomenon something that is intrinsically continuous.

- A deeper criticism: we need to distinguish between validity within-the-model and truth in the world.

Let’s start with a simple statistical model (not forgetting that it’s a model···)

- We imagine a process, and then we imagine ways in which we might misunderstand that process.

- Of course that supposes that we understand it on a deeper level.

- So we always have to be prepared for surprises.

- There are small surprises, and there are big surprises..

Using observables to estimate unobservables (I)

- Return to Equation (1)

$$(1) \qquad Y_i = \alpha + \beta . W_i + \varepsilon_i$$

- If this represents the true underlying data generating process (note this implies that *β* is the same for all *i*), and

- If error terms are distributed identically and independently of the *W*,

- Then we can use the characteristics of the observables to estimate the true causal effect

Using observables to estimate unobservables (II)

- To see this, note that

$$\frac{\sum_i Y_i(W_i = 1)}{N_1} \to \alpha + \beta \quad \text{as } N_1 \to \infty$$

$$\frac{\sum_j Y_j(W_j = 0)}{N_0} \to \alpha \quad \text{as } N_0 \to \infty$$

so that

$$\frac{\sum_i Y_i(W_i = 1)}{N_1} - \frac{\sum_j Y_j(W_j = 0)}{N_0} \to \beta$$

Using observables to estimate unobservables (III)

- We may be able to make do with weaker assumptions
- For instance, the *β* may vary in systematic ways across individuals (so long as these are independent of the *W*); see Wooldridge 4.3.3
- But in general, we can't be sure that the error terms in equation 1 will be independent of the *W*
- More commonly the underlying process will be something like

$$Y_i = \alpha + \beta.W_i + \gamma.Z_i + \varepsilon_i \qquad (2)$$

Using observables to estimate unobservables (IV)

- Note that the *Z* variables could be observed or unobserved by the researcher
- If they are all observed, the problem is simple. These are Control Variables
- What if some of the *Z* are unobserved? Three solutions:
  - Randomization, so that the Z can be treated as though they were error terms in equation 1
  - Instrumental variables, in which a pseudo-randomization is used to undertake a partial analysis of the causal effect of *W*
  - Panel data, in which the *Z* (or at least their unobservable components) are assumed to vary across individuals but be constant over time

The importance of remembering control variables

- An example: many studies (since as early as 1929) have compared breastfed children with those who were bottle-fed, and many concluded that breastfeeding led to a rise in children's IQ
- Only much later did it occur to researchers to control for the IQ of the mother
- Der et al. *BMJ* 2006 showed that much of the apparent correlation between breastfeeding and IQ shown by earlier studies is in fact due to the fact that higher-IQ mothers are more likely to breastfeed – though leaves open the question whether there might also be a causal effect

Some issues raised by the use of control variables

- Some control variables are more aggregated than the individual-level observations typical of the dependent variable
- For instance, individuals may belong to groups, and there may be a common group error term
- Then calculation of standard errors has to take into account this reduced number of degrees of freedom – by the use of *clustering* techniques (we shall look at these in more detail later in the course)
- Spatial effects may also matter when we cannot control explicitly for regions but we know that observations within a region will tend to be correlated – we need to recalculate standard errors to allow for this diminished degree of variability

What happens if some Z variables are unobserved?

- The rest of this course will be concerned with three main ways of dealing with this problem:

- Randomized controlled trials are a way of using chance to ensure that the expected influence of unobserved Z variables is small, and approaches zero in large samples

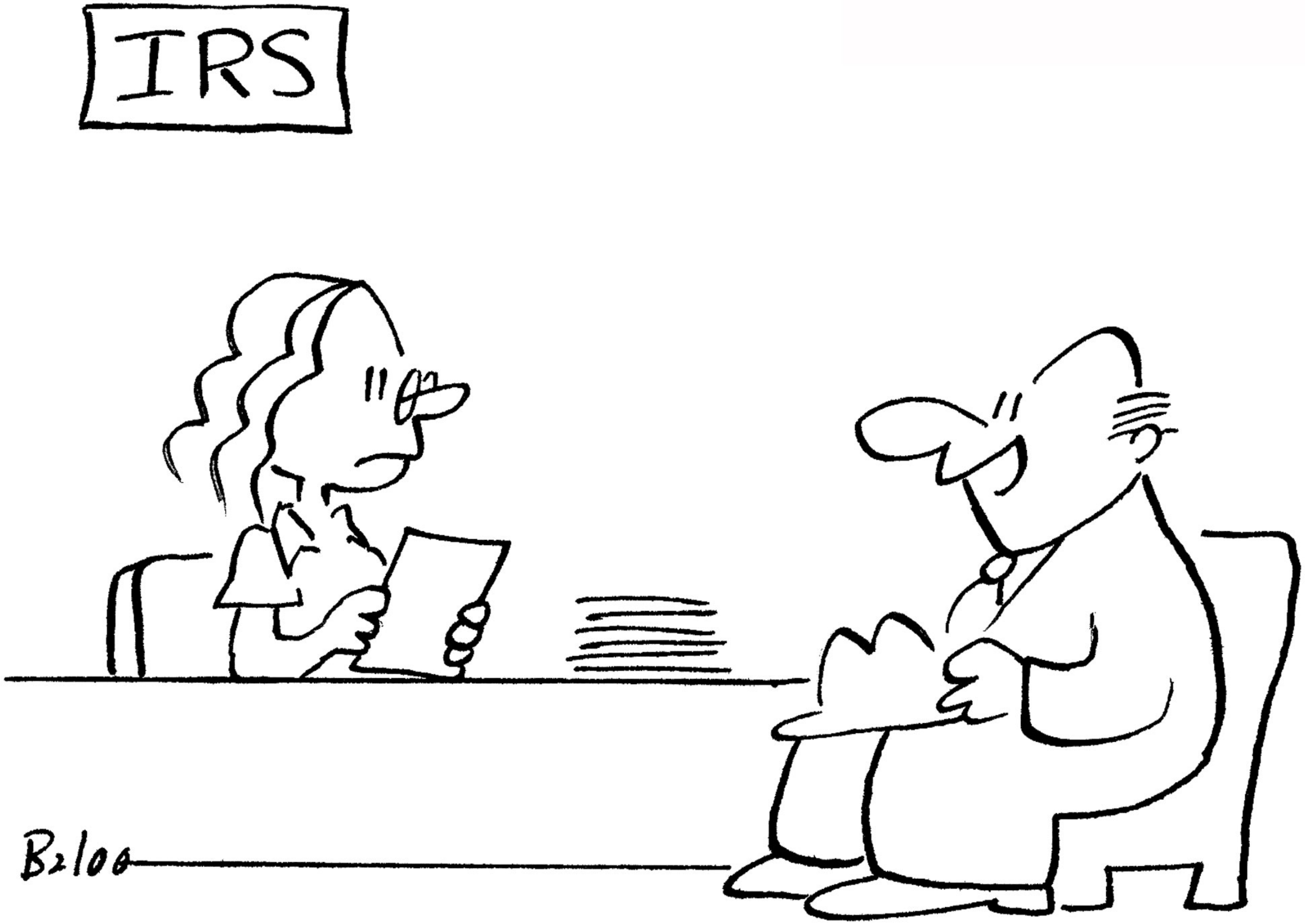
- Instrumental variables are a way of estimating causal effects of a subset of the variation in W – that subset being predicted by an exogenous variable called an instrument

- Panel data are a way of using variation over time to exclude the impact of those unobserved Z characteristics that are time-invariant

Randomization: the experimental approach

- The principle of randomization is simple: ensure that the W_i are chosen randomly so that correlation with the error terms is ruled out by construction

- In practice it is not always so easy. There may still be some tendency for correlation if
  - Sample sizes are small (eg because the experiment is expensive to run)
  - The population for the trial is chosen by the experimenter based on priors about its suitability for the treatment – then the trial population may not be representative of the wider population
  - Participation in the trial is voluntary and not all potential subjects want to participate – then there may be a difference between those who select into participation and those who do not (like the previous point, this does not cause correlation with the error term in the sample, but it does cause a bias in the parameter estimates of the wider population).



"About this new tax plan — I'd like to volunteer to be in the control group."

Placebo Effects (I)

- These are a special case of a much more general problem: uncertainties about the definition or measurement of *W*.
- Suppose the intervention involves more than one component of *W*, and the true causal process is something like this:

$$(3) \qquad Y_i = \alpha + \beta^1.W_i^1 + \beta^2.W_i^2 + \gamma.Z_i + \varepsilon_i$$

- Suppose intervention 1 is just giving people a pill, and intervention 2 is putting a specific pharmaceutical product in the pill
- Then the estimator $\hat{\beta} = \hat{\beta}^1 + \hat{\beta}^2$ will give an exaggerated estimate of the true causal impact of the pharmaceutical product

"He was unhappy to learn that I had prescribed a placebo but when I told him it was an extra-strength placebo, he was pleased."

Placebo Effects (II)

- There is evidence that in some branches of clinical medicine, placebo effects are very large
- Psychopharmacological treatments (eg anti-depressants) are a very striking example (see Kirsch 2010)
- There is even something called the “enhanced placebo effect” – a positive impact on mood of side-effects, because subjects interpret this to mean that they are in the treatment, not the placebo group!
- Pain medication shows strong placebo effects (Goldacre 2008, ch. 5)
- Can you think of analogies in political and economics contexts?

Placebo Effects (III)

- Uncertainties about the definition of *W* do not necessarily lead to exaggerated estimates. They can be downward-biased too. Suppose there is error in the measurement of *W*, so that:

$$(4) \qquad Y_i = \alpha + \beta.w_i + \gamma.Z_i + \varepsilon_i$$

$$\text{where } w_i = W_i + e_i \text{ and } \mathrm{Cov}(W_i, \varepsilon_i) = 0$$

- Then the estimator $\hat{\beta}$ will give an dampened estimate of the true causal impact of the pharmaceutical product, because

$$\mathrm{Cov}(w_i, \varepsilon_i) > 0$$

Externalities

- Suppose that in fact the true causal process is

$$(4) \qquad Y_i = \alpha + \beta^1 . W_i^1 + \sum_{j \neq i} \beta^2 . W_j^1 + \gamma . Z_i + \varepsilon_i$$

But that the researcher mistakenly estimates equation (1)

- Then the estimator $\hat{\beta}$ will give a downward biased estimate of the true causal effect, for two reasons:
  - The error term in the observed equation for individual *i* will contain the externalities from individuals *j* – which may be correlated with *W*, thereby biasing the estimate of β^1
  - The true causal effect should include the externalities of *i* on all the *j*, so the estimator β^1 is a downward-biased measure of the true causal effect (if the externalities are positive)

Generalizability of findings (I)

- Much concern has been expressed about whether finding from randomized controlled trials can be expected to hold when the intervention is rolled out to a wider population

- One possible reason for skepticism is selection effects (by the experimenter or by the subjects) - In particular, we need to distinguish “Treatment” from “Intention to Treat”.

- Another reason is problems in identifying the true nature of the intervention – was it really *W* or some other factor (eg the quality or the enthusiasm of the trial’s staff)?

- In particular, interventions can have unexpected consequences that need to be considered in the proper description of the treatment

An example of unexpected consquences – from the 2nd century AD!

- From the essay “Alexander the False Prophet,” by Lucian, Greek rhetorician: “There was one oracle, also an autophone, which he [[Alexander]] dispatched to all the nations during the pestilence [the terrible plague which swept the whole Empire about A.D. 165]: ‘Phoebus, the god unshorn, keepeth off plague’s nebulous onset.’

- This verse was to be seen everywhere written over doorways as a charm against the plague; but in most cases it had the contrary result. By some chance it was particularly the houses on which the verse was inscribed that were depopulated! Do not suppose me to mean that they were stricken on account of the verse—by some chance or other it turned out that way, and perhaps, too, people neglected precautions because of their confidence in the line and lived too carelessly, giving the oracle no assistance against the disease because they were going to have the syllables to defend them and ‘unshorn Phoebus’ to drive away the plague with his arrows!”

Generalizability of findings (II)

- Deaton (2010) emphasizes difficulty of identifying treatment effects without understanding of the mechanisms being investigated.

- Therefore applying experimental results to new contexts is risky unless we can know that the new contexts replicate the mechanisms

- In many cases it's unlikely there is a uniform treatment effect across all subjects. (Consider medical examples of Yervoy, Erbitux and Herceptin, cancer treatments that work only on a subset of treated subjects). Thus generalizing requires considering which characteristics of subjects are relevant for affecting magnitude of treatment effects.

- Even if other characteristics are orthogonal to treatment, calculation of standard errors will be affected.

Strengths and weaknesses of the experimental approach

- Permits direct methods of controlling for potential differences in groups that are and are not affected by an intervention – can be much more reliable than purely statistical methods
- Can allow for refinement of hypothesis testing – if an experiment reveals unexpected results it can be modified relatively quickly to probe these results further
- However, many important causal questions are not susceptible to experimental treatment – does democracy help economic growth for example?
- And experimental interventions are not always representative of ones that can realistically be undertaken given political constraints

"Do a double-blind test. Give the new drug to rich patients and a placebo to the poor. No sense getting their hopes up. They couldn't afford it even if it works."