Instrumental variables
An intuitive guide

Pierre-Louis Vézina

University of Oxford

June 8, 2014

The quest for causal evidence

- What causes growth?
- What causes good schooling?
- What causes productive agriculture?

A causal relationship is useful for making predictions about the effect of policies; it tells us what would happen in a counterfactual world

The quest for causal evidence

The causal question: Does chocolate consumption increase your chances of winning the Nobel Prize?

The quest for causal evidence

Evidence from the New England Journal of Medicine

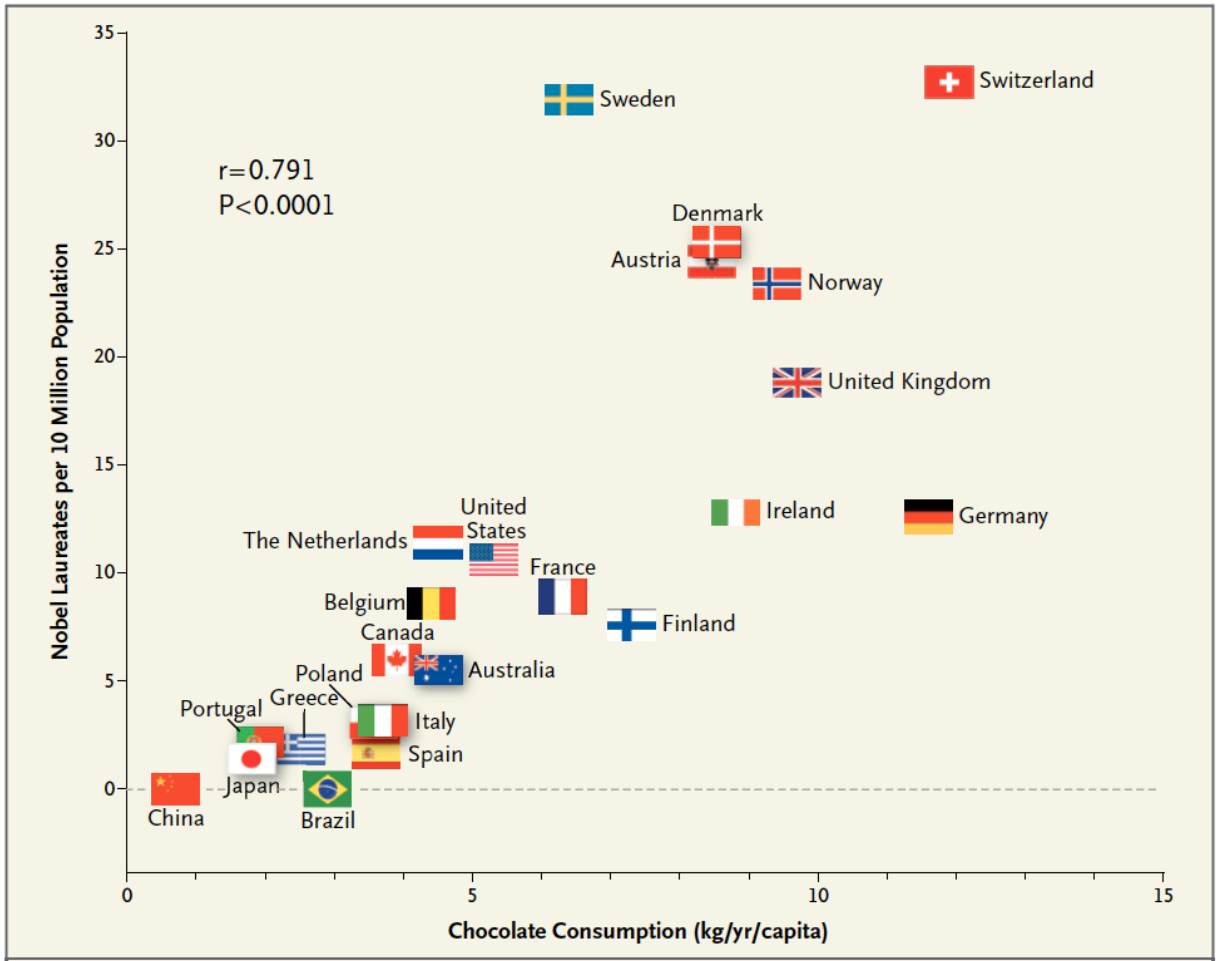

Figure 1. Correlation between Countries' Annual Per Capita Chocolate Consumption and the Number of Nobel Laureates per 10 Million Population.

Causality vs. correlation

- Is this convincing evidence chocolate causes Nobel Prizes? Why not?
- Rich countries probably consume more chocolate AND produce top scientists
- Despite the possibility to include controls, such as GDP per capita, in OLS regressions, there is often an unobservable that may drive the correlation

Causality vs. correlation

- What makes a correlation questionable?
  - Unobservables driving the relationship
  - Reverse causality, e.g. more police leads to more crime

Causality vs. correlation

- What can we do to show chocolate causes Nobel Prizes?
- Distribute chocolate bars randomly to some scientists in various cities across countries and see if those that eat chocolate earn more Nobel Prizes
- This is hard to do, at least on a big scale, and it takes years

Causality vs. correlation

- What else can we do?
  - We can look for such an *accidental* distribution in history
  - Start by asking yourself: Why do some countries consume chocolate?
    - Do they grow cocoa? Produce milk? Grow sugarcane?
    - Did they use to trade with the Aztecs?
    - Do they have a colonial link to cocoa-growing countries?
    - Are any of these explanations completely unrelated to Nobel Prizes?
  - If so, such factors can be used as "instrumental variables" (IVs) to establish causality
  - The variation of chocolate consumption caused by the IV will not be correlated with the unobservables causing Nobel Prizes

Finding an instrument

- A valid instrument needs to satisfy three conditions:
  - It must be strongly correlated with your endogenous variable, X
  - It must be as good as random, i.e. uncorrelated with the unobservables
  - It must be excludable, i.e. it affects Y only through X

Finding an instrument

- Common sources of instruments include:
  - Nature: Geography, weather, biology in which a truly random source of variation influences X
  - History: Things determined a long time ago which no longer plausibly influence Y
  - Formal or informal rules that influence the assignment of X in a way unrelated to Y
- Above all, finding a good IV is based on deep knowledge of the processes shaping X and Y
  - Fancy econometrics will not help you

Using an instrument

- Z is the IV, X the policy variable, and Y the outcome
- $Z \Rightarrow X \Rightarrow Y$
- Three IV estimation methods:

Using an instrument

- Z is the IV, X the policy variable, and Y the outcome
- $Z \Rightarrow X \Rightarrow Y$
- Three IV estimation methods:
- 1. Two-Stage Least Squares (2SLS): First, use OLS to regress X on Z and get $\hat{X}$, then use OLS to regress Y on $\hat{X}$ to get β_{IV}

Using an instrument

- Z is the IV, X the policy variable, and Y the outcome
- $Z \Rightarrow X \Rightarrow Y$
- Three IV estimation methods:
- 2. Ratio of Coefficients: Regress X on Z to get β_{XZ} and regress Y on Z to get β_{YZ}, and the IV estimate $\beta_{IV} = \beta_{YZ}/\beta_{XZ}$. If X is a binary indicator variable, this method is known as the Wald estimator.

Using an instrument

- Z is the IV, X the policy variable, and Y the outcome
- $Z \Rightarrow X \Rightarrow Y$
- Three IV estimation methods:
- 3. The Control Function Approach: Use OLS to regress X on Z and get estimated errors $\hat{v}$, then use OLS to regress Y on X and $\hat{v}$ to get β_{IV}

Examples of Instruments

- Do institutions affect economic performance?
  - 'The Colonial Origins of Development' by Acemoglu, Johnson, and Robinson
  - Use settler mortality rate as an IV
  - Settlements $\Rightarrow$ Early Institutions $\Rightarrow$ Current Institutions $\Rightarrow$ Current economic performance
  - Exclusion restriction: Mortality rates of European settlers more than 100 years ago have no effect on current income per capita other than via institutions

Examples of Instruments

- How has colonialism affected modern income?
  - 'Colonialism and Modern Income: Islands as Natural Experiments' by Feyrer and Sacerdote
  - The early colonialists went where their sails took them before steamships
  - Some islands were colonised early, some late, for reasons that had little to do with any attractions the islands might offer
  - Use winds and currents as IV for period of colonisation of 80 islands
  - Exclusion restriction: Have winds and currents not affected prosperity through other channels than colonialism?

Examples of Instruments

- Do immigrants promote international trade?
  - 'Migrant Networks and Trade: The Vietnamese Boat People as a Natural Experiment' by Parsons and Vézina
  - Vietnamese refugees were distributed across US States in 1975
  - Instrument 1995 Vietnamese migrant stocks with 1975 refugee allocation
  - Exclusion restriction: Refugees affect trade with Vietnam only through Vietnamese networks
  - Is it correlated with the unobservables? Was the allocation random?

Examples of Instruments

- Do economic conditions affect civil war?
  - 'Economic Shocks and Civil Conflict' by Miguel, Satyanath, and Sergenti
  - Weather strongly predicts GDP in countries that rely on rain for agriculture and are prone to drought
  - Instrument GDP with rainfall
  - Exclusion restriction: Weather should not affect likelihood of conflict except through its influence on economic growth

Pitfalls of instruments

- Instruments can always be taken down
  - Be ready to face the 'instrument police'
  - You have to be more convincing than your critics
- Do not confuse the cute and the rigorous
  - Your aim is to provide causal evidence, not to show off ingenuity
- The IV is not your variable of interest, leave it out of your policy discussion