

# Verejné financie II.

## Seminár VII - **Empirická analýza**

Matej Lorko  
matej.lorko@euba.sk

Slidy a materiály: [www.lorko.sk](http://www.lorko.sk)

### Readings:

- Gruber, J. (2005). Public finance and public policy. Macmillan.
- Congdon, W. J., Kling, J. R., & Mullainathan, S. (2011). Policy and choice: Public finance through the lens of behavioral economics. Brookings Institution Press.
- Sunstein, C. R. (2020). Behavioral science and public policy. Cambridge University Press.

- Je známym faktom, že vysoko inteligentné ženy majú v priemere menej inteligentných manželov. Prečo?

# Empirická analýza

- Základným problémom pri otázkach verejných politík je vyhodnotenie kauzálneho vzťahu medzi politikou a jej efektom
  - Spôsobí zvýšenie prídavkov na deti vyššiu nezamestnanosť u slobodných matiek?
  - Spôsobuje viac znečistené ovzdušie viac zdravotných následkov?
  - atď.
- V rámci tejto prednášky budeme diskutovať o viacerých spôsoboch oddelenia kauzálnosti od obyčajnej korelácie. Zlatým štandardom sú v tomto smere kontrolované randomizované experimenty (randomized controlled trial, RCT), ktoré pomocou náhodného rozdeľovania do intervenčných a kontrolných skupín minimalizujú systematické skreslenia v dátach. Nanešťastie, RCT nie je možné realizovať pre všetky otázky, preto si musíme vypomáhať alternatívnymi metódami, napr. analýzou časových radov, regresnou analýzou a kvázi experimentálnou analýzou.

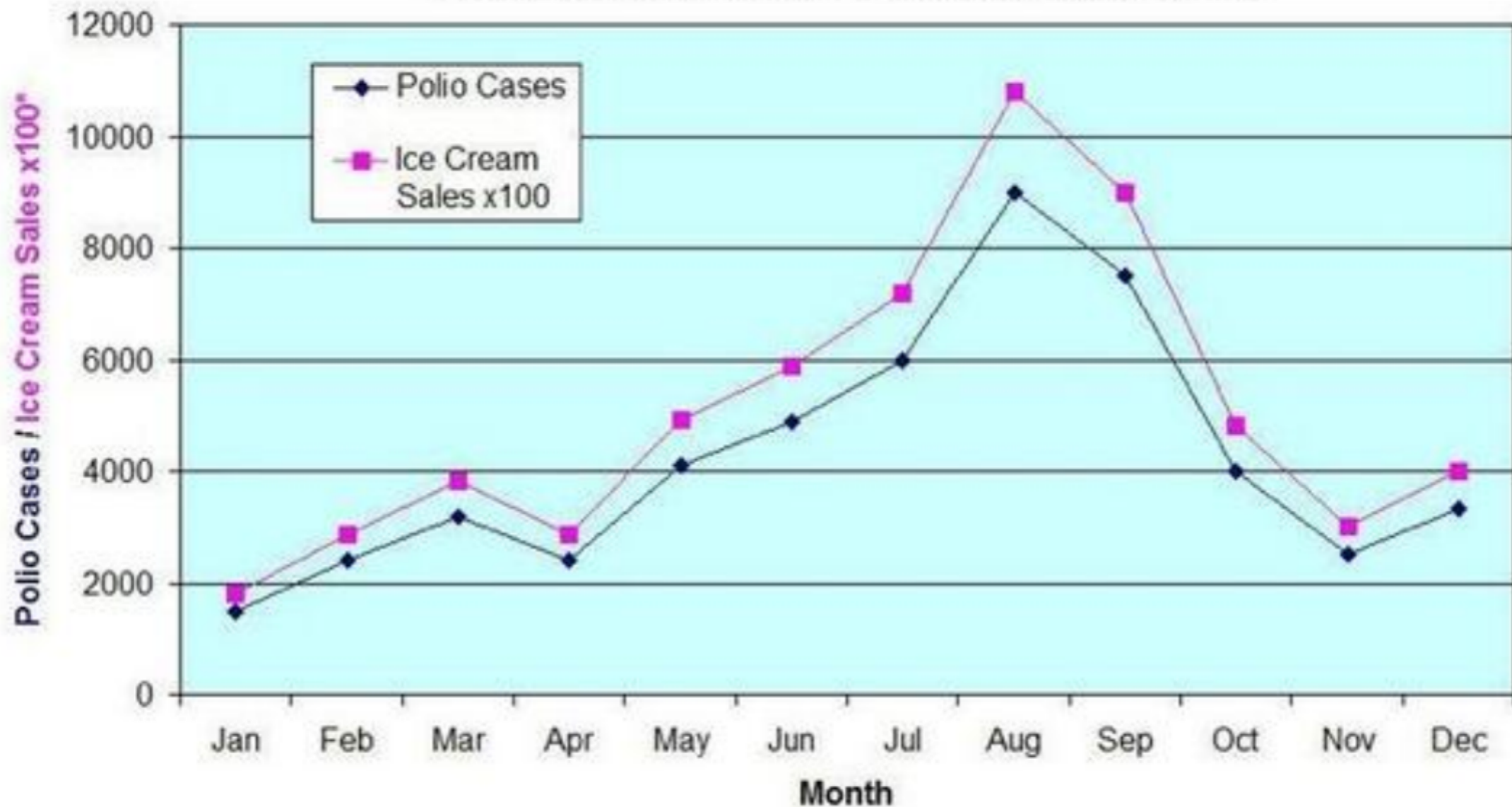
# Empirická analýza

- **Empirické verejné financie:** využitie dát a štatistických metód na meranie dopadov verejných politík na jednotlivcov a trhy.
- Fundamentálnym problémom každej empirickej štúdie v ekonómii je oddelenie kauzálnosti od korelácie
- O dvoch premenných hovoríme ako o korelovaných, ak sa ich hodnoty menia súčasne. Na to, aby ich vzťah bol aj kauzálny, však pohyb jednej premennej musí spôsobovať pohyb druhej. Tým sa dostávame k identifikačnému problému. Ak sú dve premenné korelované, ako zistíme či jedna ovplyvňuje druhú?

# Korelácia vs. kauzálnosť

## The Real Cause of Polio!

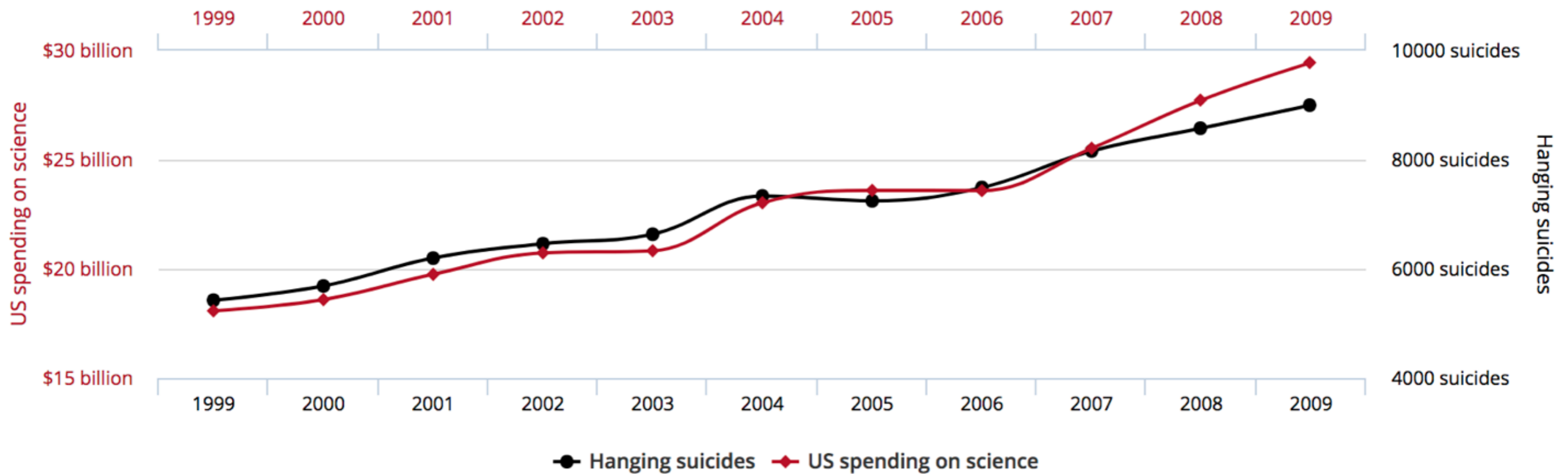
Polio Rates / Ice Cream Sales 1949



# Korelácia vs. kauzálnosť

US spending on science, space, and technology  
correlates with  
Suicides by hanging, strangulation and suffocation

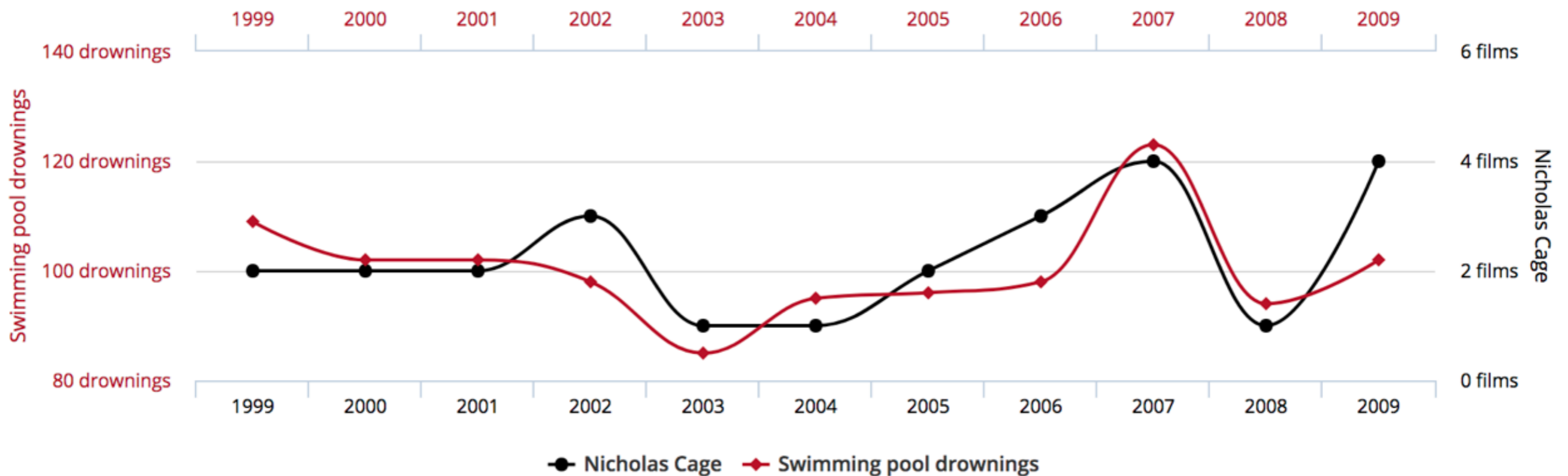
Correlation: 99.79% ( $r=0.99789126$ )



# Korelácia vs. kauzálnosť

Number of people who drowned by falling into a pool  
correlates with  
Films Nicolas Cage appeared in

Correlation: 66.6% (r=0.666004)



# Korelácia vs. kauzálnosť

- Cholera v Rusku
- SAT kurzy a výsledky testov
- Kojenie a podvýživa
- ERT
- Zostavy hokejových U20 reprezentácií



# Korelácia vs. kauzálnosť

- Postup analýzy
  - Krok 1: Zdokumentovanie korelácie, teda, či sa dáta dvoch premenných pohybujú súčasne.
  - Krok 2: Vyšetrenie, či pohyby jednej premennej spôsobujú pohyby druhej.
    - Pre akúkoľvek koreláciu medzi dvoma premennými A a B môžu existovať 3 vysvetlenia:
      - A ovplyňuje B
      - B ovplyvňuje A
      - Existuje iný faktor, ktorý ovplyvňuje zároveň A aj B

# Posudzovanie kauzálnosti

- SAT:
  - A  $\rightarrow$  B: SAT kurzy zhoršujú pripravenosť na SAT skúšku.
  - B  $\rightarrow$  A: Na SAT kurzy chodia najslabší študenti, snažiaci sa dobehnúť lepších
  - C  $\rightarrow$  A,B: Ľudia, ktorí sú vo všeobecnosti nervózni/stresujú, sú tí, ktorí pravdepodobne radšej pôjdu na kurz, ale zároveň kvôli ich stresovaniu nezvládajú dobre testy.
- Kojenie:
  - A  $\rightarrow$  B: Dlhšie kojenie škodí zdraviu detí.
  - B  $\rightarrow$  A: Deti, ktoré sú nezdravé, sú kojené dlhšie.
  - C  $\rightarrow$  A,B: Matky z chudobných pomerov koja dlhšie, keďže je to najlacnejšia forma výživy, ale zároveň sú deti kvôli chudobe nezdravé.

- Skupine depresívnych detí bol denne podávaný 1 liter Coca Coly. Po niekoľkých mesiacoch sa zistilo, že deti sú na tom lepšie. Mali by sme začať liečiť depresiu Coca Colou?

# Zlatý štandard pre posudzovanie kauzálnosti: RCT

- Z korelácie teda ešte priamo nemusí vyplývať kauzálnosť. Inými slovami, to, že sú dve premenné korelované, nám neposkytuje žiadnu schopnosť predikcie, ak nevieme, či je medzi nimi aj kauzálny vzťah. Ako kauzálnosť odhaliť? Najlepším riešením sú RCT.
- RCT (randomized control trial) - experiment designovaný priamo na odhalenie kauzálnosti a to tak, že skupina jednotlivcov je náhodne (hodom mince) rozdelená do intervenčnej skupiny, na ktorej je testovaná určitá zmena alebo intervencia a do kontrolnej skupiny, ktorá intervenciu nedostane.

# Zlatý štandard pre posudzovanie kauzálnosti: RCT

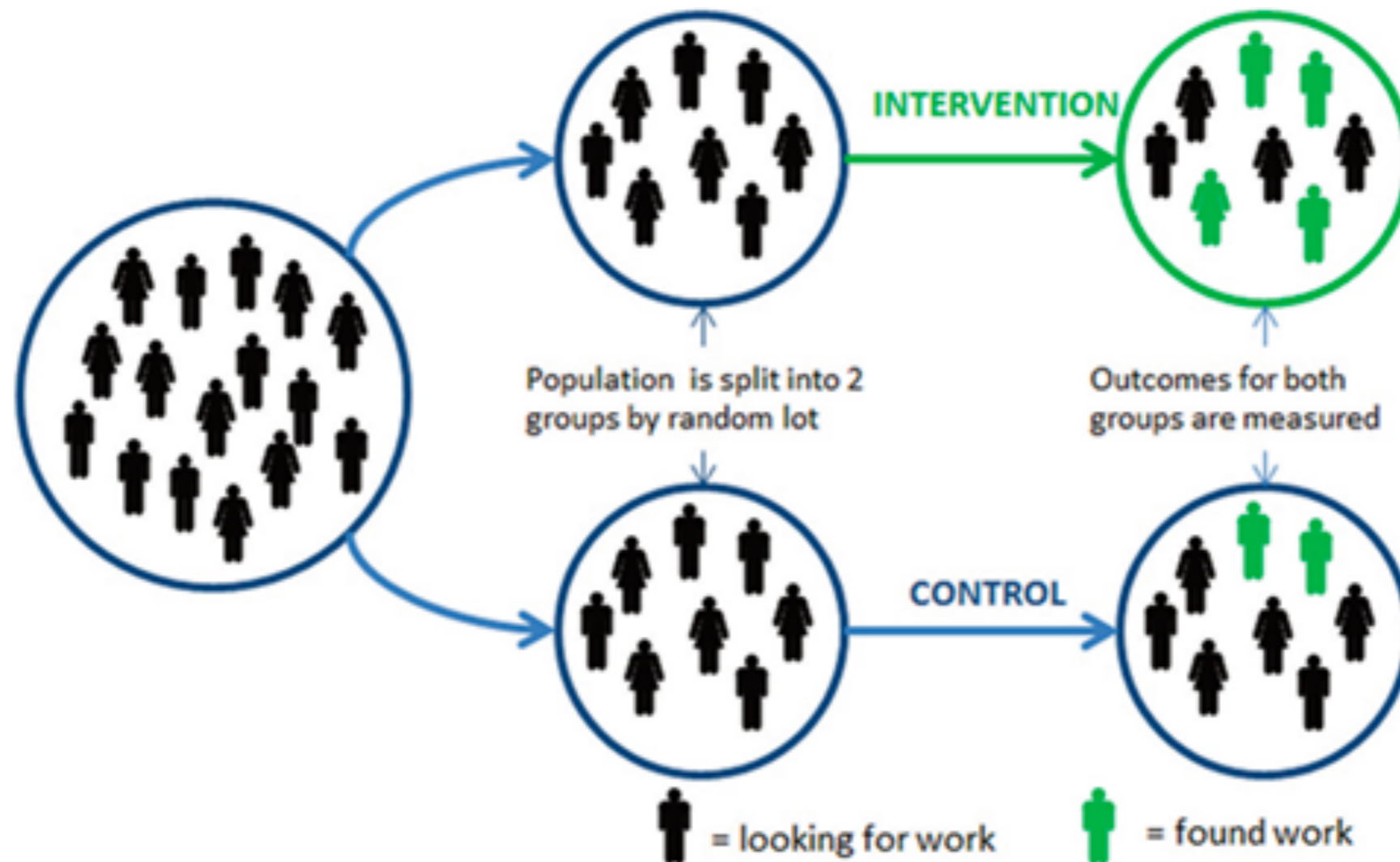


Figure 1. The basic design of a randomised controlled trial (RCT), illustrated with a test of a new 'back to work' programme.

# Problém skreslenia (bias)

- Každú analýzu empirickej metodológie by sme mali začať jednoduchou otázkou - líšia sa intervenčná a kontrolná skupina aj z nejakého iného pohľadu, ako iba intervenciou?
- Rozdiely medzi intervenčnou a kontrolnou skupinou, ktoré nemajú súvis so samotnou intervenciou sú fundamentálnym problémom pri určovaní kauzálnosti z korelácie. Tieto rozdiely nazývame “skreslenia” (anglicky bias). Skreslenie je akýkoľvek zdroj rozdielov medzi intervenčnou a kontrolnou skupinou, ktorý je s intervenciou korelovaný, ale nie ňou spôsobený.
- Keďže sú pri RCT jednotlivci rozdeľovaní do skupín náhodne, nemali by medzi skupinami (pri dostatočne veľkej vzorke jednotlivcov - zákon veľkých čísel) takéto systematické rozdiely spôsobujúce skreslenia existovať. Tým pádom by mala byť jedinou odlišnosťou medzi skupinami práve intervencia - a to je dôvod, prečo sú RCT zlatým štandardom pre určovanie kauzálnych efektov intervencií.

# RCT nedokáže všetko...

- Pri niektorých otázkach je však použitie RCT v podstate nemožné...
  - Majú veteráni menšie mzdy, kvôli tomu že strávili čas na vojne?
  - Zvyšuje vysokoškolský titul mzdu?
  - Aký je sklon krivky dopytu?
  - Ako ovplyvňuje minimálna mzda zamestnanosť?
  - Má veľkosť triedy vplyv na štúdijné výsledky?
  - Aká hladina alkoholu krvi zvyšuje riziko havárie?
  - Zlepšuje prístup k informáciám efektívnosť trhov?

# Nedostatky RCT

- Pre mnohé otázky nemôžeme RCT použiť, napr. preto, že experiment by bol príliš drahý, trval príliš dlho, alebo by bol eticky nepriechodný (napr. niektoré nové medicínske procedúry)
- Navyše, aj metóda RCT má svoje nedostatky. Prvý je, že výsledky sa merajú zo vzorky jedincov, často dobrovoľníkov, ktorá môže byť trochu odlišná od charakteristík všeobecnej populácie. Napríklad môžu mať účastníci štúdie menšiu averziu k riziku, alebo môžu byť vážne (zúfalo) chorí. Tým pádom sa môže stať, že výsledok RCT nebude validný pre priemerného človeka z populácie.
- Druhým problémom je prirodzený úbytok (attrition). Jednotlivci môžu experiment opustiť skôr, ako sa skončí. Ak by experiment opúšťali náhodne, nebol by to problém, lebo by vzorky ostali vyvážené. Predpokladajme však, že intervencia má pozitívne efekty na polovicu intervenčnej skupiny a negatívne na druhú polovicu, a že tí, na ktorých pôsobí intervencia negatívne, experiment opustia. Keďže v experimente ostanú iba tí, ktorým intervencia pomáha, môže sa stať, že na konci štúdie budeme nesprávne konštatovať, že intervencia má pozitívny efekt.
- prirodzený úbytok (attrition): zmenšenie vzorky v priebehu experimentu. Toto zmenšenie, ak nie je náhodné, môže viesť k skresleniu výsledkov.

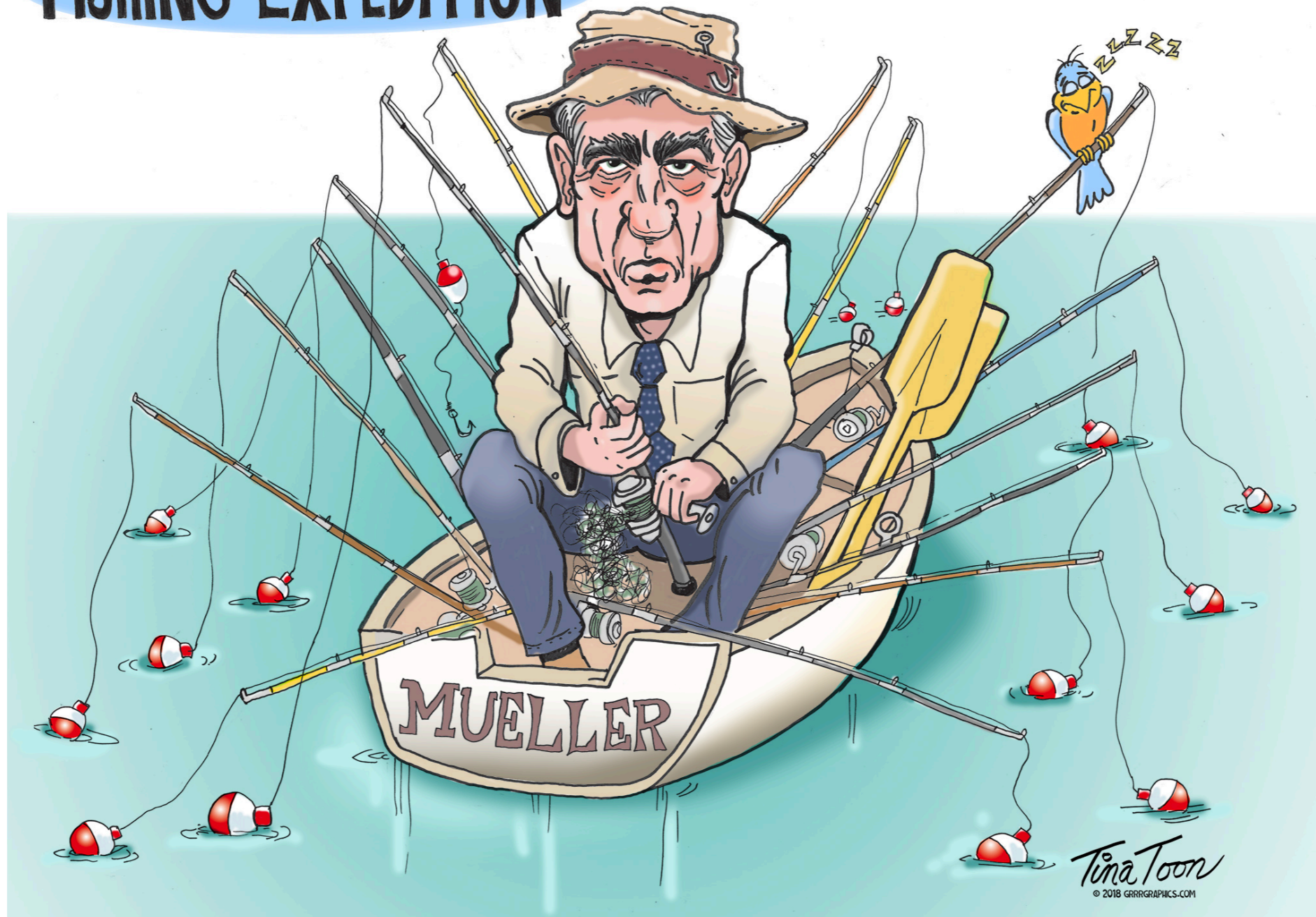


# Určovanie kauzálnosti z neexperimentálnych dát

- Ak nie sú k dispozícii dáta z RCT, musíme sa spoľahnúť na klasické dáta z pozorovaní (observácií)
- Observačné dáta: dáta vygenerované správaním jednotlivcov v “reálnom svete”, teda mimo zámerne nadizajnovných experimentov
- Napríklad pri nových liekoch môžeme analyzovať, namiesto RCT dát, dáta o tom, kto liek užíval a aké boli výsledky
- Samozrejme, observačné dáta sú skreslené (napr. vitamíny užívajú iba špecifické typy ľudí, nie všetci), našťastie však máme k dispozícii niekoľko nástrojov, ktorými vieme so skreslením pracovať tak, aby bolo možné (s určitou nutnou dávkou rezervy) určovať kauzálnosti aj v takýchto, neexperimentálnych dátach

# Žiadna hypotéza, žiaden výsledok

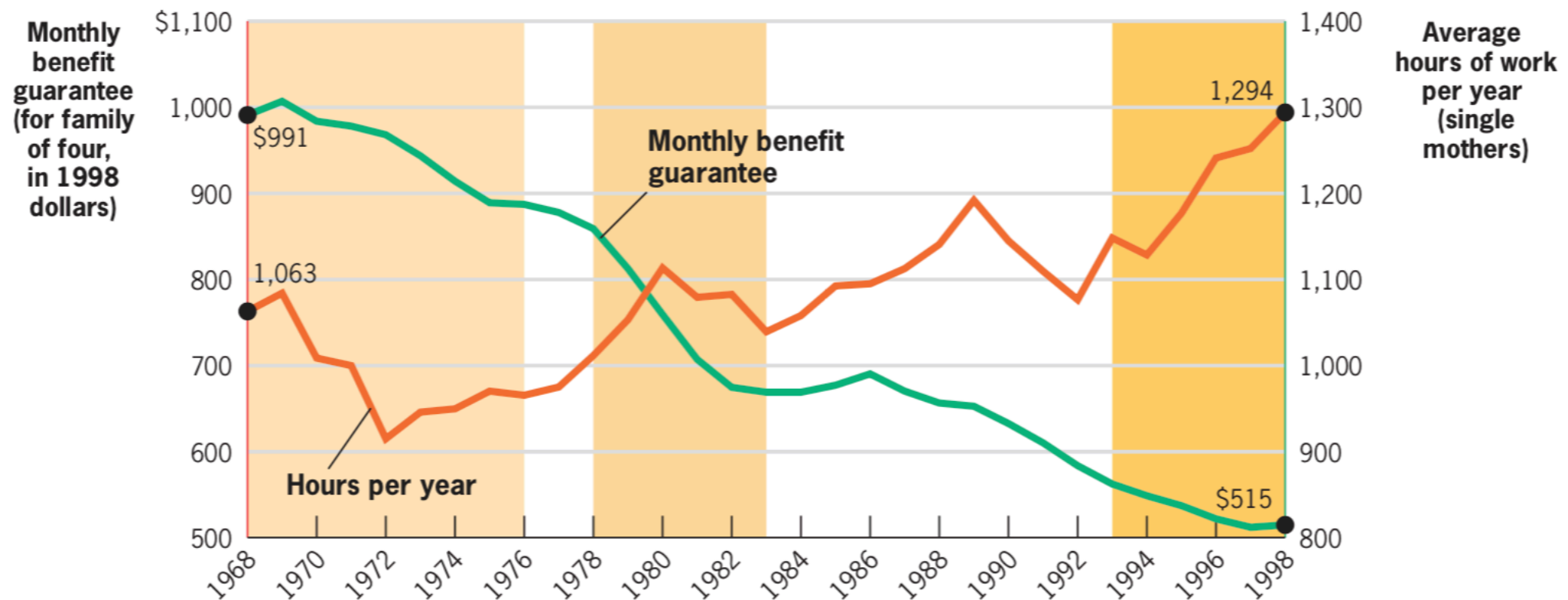
FISHING EXPEDITION



# Analýza časových radov

- Pri určovaní kauzálnosti z observačných dát sa často využíva analýza časových radov, teda analýza zmien dvoch premenných v čase

■ FIGURE 3-1



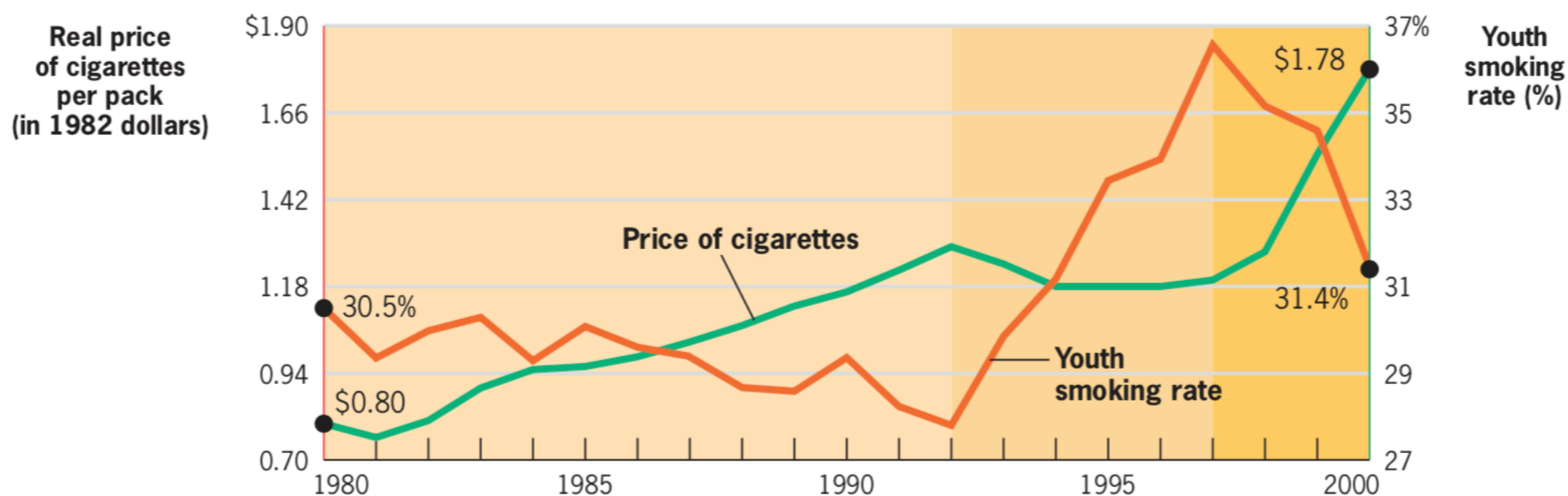
**Average Benefit Guarantee and Single Mother Labor Supply, 1968–1998** • The left-hand vertical axis shows the monthly benefit guarantee under cash welfare, which falls from \$991 in 1968 to \$515 in 1998. The right-hand vertical axis shows average hours of work per year for single mothers, which rises from 1,063 in 1968 to 1,294 in 1998. Over this entire 30-year period, there is a strong negative correlation between the average benefit guarantee and the level of labor supply of single mothers, but there is not a very strong relationship within subperiods of this overall time span.

# Analýza časových radov

- Nedostatky
  - Ani veľmi silná korelácia premenných v čase ešte nemusí znamenať kauzálny vzťah, keďže v hre môžu byť aj iné faktory, ktoré analýza nemusí zachytiť. Takéto faktory môžu spôsobovať skreslenia, ak sú tiež korelované so sledovanými premennými.
- Kedy je užitočná?
  - Hlavne ak v dátach nájdeme výrazné skoky, ktoré nie sú spojené s tretími faktormi. Naopak, časové rady nám nedokážu veľa povedať, ak nájdeme dlhé a pomerne nevýrazné trendy.

# Analýza časových radov

■ FIGURE 3-2



**Real Cigarette Prices and Youth Smoking, 1980–2000** • The left-hand vertical axis shows the real price of cigarettes per pack, which rises from \$0.80 in 1980 to \$1.78 in 2000. The right-hand vertical axis shows the youth smoking rate (the share of high school seniors who smoke at least once a month), which fell from 1980 to 1992, rose sharply to 1997, and then fell again in 2000 to roughly its 1980 level. There is a striking negative correspondence between price and youth smoking within subperiods of this era.

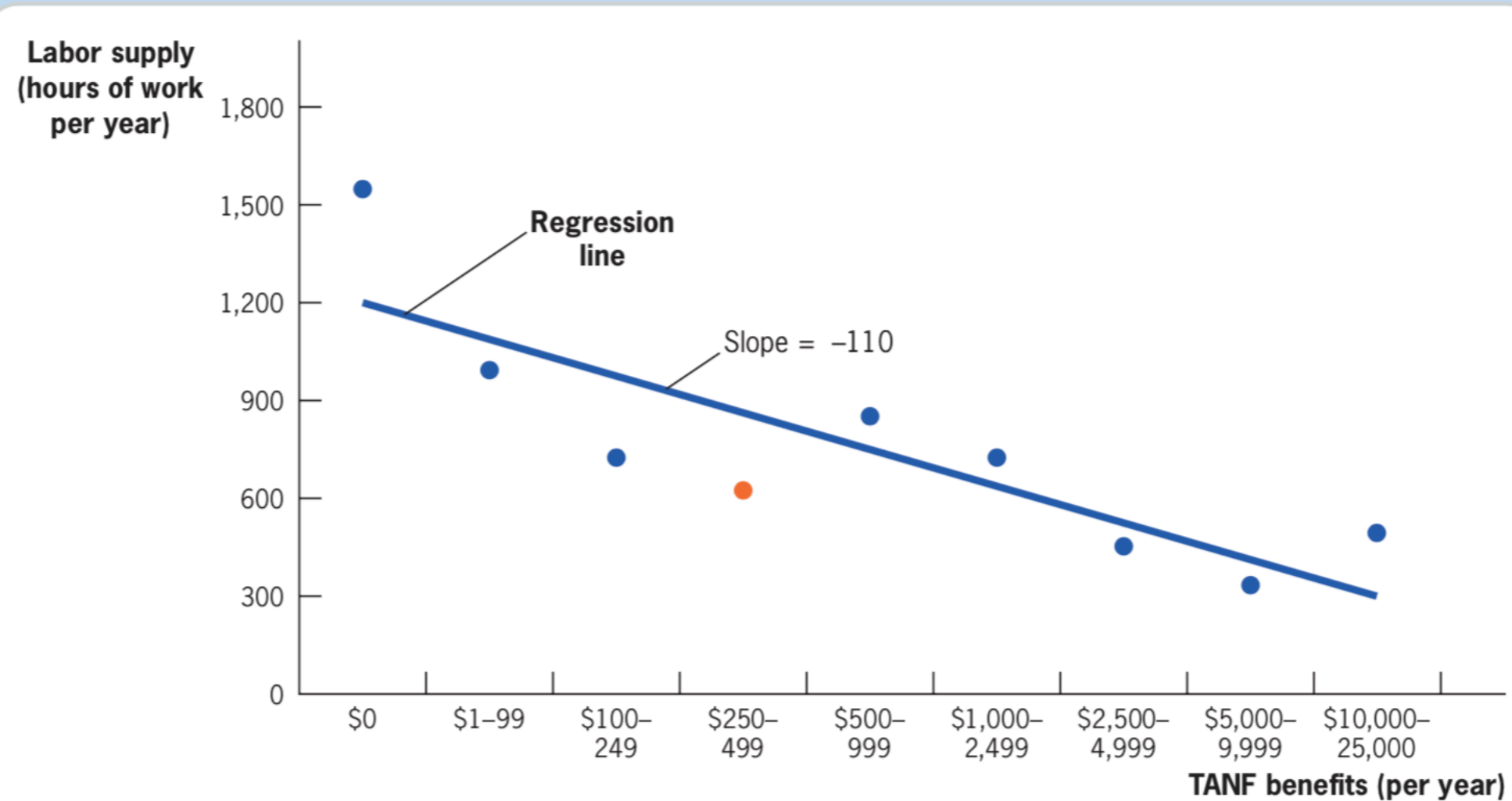
Source: Calculations based on data on smoking from Monitoring the Future survey and on tobacco prices from the Tobacco Institute.

# Prierezová regresná analýza

- Ďalším postupom pre identifikovanie kauzálnych efektov je prierezová regresná analýza, ktorá vyhodnocuje vzťah dvoch premenných, pričom drží ostatné faktory konštantné. “Prierezová” znamená porovnanie množstva jednotlivcov v jednom časovom bode (teda nie vo viacerých obdobiach ako to robia časové rady)
- Regresná analýza popisuje (a kvantifikuje) vzťah medzi premennou, ktorej vývoj sa snažíme objasniť (závislá premenná) a súborom premenných, o ktorých sa domnievame, že by ju vysvetliť mohli (nezávislé premenné)
- Odhad takéhoto vzťahu zobrazuje regresná priamka. Keďže priamka, ktorá by prešla všetkým bodom v dátovom súbore neexistuje (museli by sme nakresliť krivku), lineárna regresia hľadá priamku, ktorá je k bodom čo najbližšie
- Technicky priamka identifikuje súčet najmenších štvorcov, teda kvadrátov vzdialenosti jednotlivých bodov od priamky. Táto technika je tým pádom citlivá na body, ktoré sú od nej príliš ďaleko (outliers), keďže im dáva veľmi vysokú váhu a teda vplyv. Niekedy je preto nutné zvážiť aj postupy, ktoré budú k vzdialeným pozorovaniam menej citlivé.

# Prierezová regresná analýza

■ FIGURE 3-4



**TANF Benefit Income and Labor Supply of Single Mothers, Using CPS Data** • Using data from the CPS, we group single mothers by the amount of TANF income they have. Those who are receiving the lowest level of TANF income are the ones providing the highest number of work hours.

Source: Calculations based on data from Current Population Survey's annual March supplements.

# Prierezová regresná analýza

- Vzťah medzi dvoma premennými, ktorý aproximujeme regresnou priamku nemusí byť kauzálny. Preto výsledky regresie neinterpretujeme slovami “x-percentný pokles/nárast premennej A spôsobuje y-percentný pokles/nárast premennej B”, ale “x-percentný pokles/nárast premennej A je spojený s y-percentný pokles/nárast premennej B”
- Regresná analýza má oproti korelácii výhodu v tom, že dokáže pracovať so skreslením, vďaka možnosti zahrnúť do nej kontrolné premenné. Kontrolné premenné vnesú do analýzy iné ako intervenčné rozdiely medzi jedincami vo vzorke, a tým pádom zostávajúcu koreláciu medzi A a B je možné o trochu spoľahlivejšie označiť za kauzálnosť.
- Použitie kontrolných premenných však nikdy nevyrieši problém skreslenia úplne, keďže väčšinou nemáme k dispozícii všetky charakteristiky, ktoré by mohli závislé premenné ovplyvňovať.



# Prierezová regresná analýza

- $y = \alpha + \beta x + e$
- kde
  - $\alpha$  = konštanta (hodnota pre  $x = 0$ )
  - $\beta$  = koeficient sklonu, vyjadrujúci zmenu  $y$  pri jednotkovej zmene  $x$
  - $e$  = odchýlka, ktorá vyjadruje rozdiel medzi skutočnou hodnotou premennej a jej hodnotou odhadovanou modelom

# Prierezová regresná analýza

Linear regression		Number of obs	=	130		
		F(12, 117)	=	3.49		
		Prob > F	=	0.0002		
		R-squared	=	0.2353		
		Root MSE	=	586.8		
abserror	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
After	-252.8127	118.7671	-2.13	0.035	-488.0247	-17.60077
Before	-413.7776	111.5208	-3.71	0.000	-634.6386	-192.9165
Detail	283.7623	180.3411	1.57	0.118	-73.3938	640.9185
dem1	-1.658151	19.99805	-0.08	0.934	-41.26323	37.94693
dem2	15.81251	111.2361	0.14	0.887	-204.4847	236.1097
dem3	69.00018	50.26583	1.37	0.172	-30.54867	168.549
dem4	49.68547	53.67132	0.93	0.356	-56.60777	155.9787
dem6	-68.4688	52.44197	-1.31	0.194	-172.3274	35.38977
riskONE	-5.618447	25.47368	-0.22	0.826	-56.06774	44.83084
timeEST	-1.795241	1.531554	-1.17	0.244	-4.828403	1.237921
timeCONF	7.61295	10.98917	0.69	0.490	-14.15053	29.37643
confidence	-26.79666	74.69199	-0.36	0.720	-174.7202	121.1269
_cons	661.6392	568.5706	1.16	0.247	-464.3849	1787.663

**Table 6:** Linear regression analysis

Dependent variable	(1) Estimate	(2) Actual Duration	(3) Estimation bias	(4) Absolute estimation error
1. Info-After Treatment	323.63** (-140.63)	-94.88 (-131.26)	380.79** (-161.68)	-252.81** (-118.77)
2. Info-Before Treatment	216.79* (-127.3)	-124.96 (-124.76)	316.48** (-150.28)	-413.78*** (-111.52)
3. Detailed Description Treatment	586.68** (-266)	-8.68 (-146.65)	526.98* (-267.93)	283.76 (-180.34)
4. Age	-5.54 (-23.4)	5.46 (-19.59)	-10.35 (-25.66)	-1.66 (-20)
5. Female	-84.79 (-153.07)	-88.01 (-96.09)	13.1 (-174.69)	15.81 (-111.24)
6. Self-reported math skill	27.8 (-79.59)	-188.03*** (-53.23)	212.58*** (-79.99)	-69 (-50.27)
7. Current degree of study	34.51 (-79.7)	1.51 (-41.54)	28.97 (-72.4)	49.69 (-53.67)
8. Employment status	52.91 (-68.78)	-10.26 (-47.06)	57.01 (-78.08)	-68.47 (-52.44)
9. Risk attitudes	45.22 (-30.23)	28.89 (-26.61)	11.06 (-35.89)	-5.62 (-25.47)
10. Time spent estimating	-2.83 (-2.37)	-1.73 (-1.26)	-0.77 (-2.65)	-1.8 (-1.53)
11. Time spent indicating confidence in estimate	10.72 (-15.37)	11.24 (-12.69)	-1.77 (-19.28)	7.61 (-10.99)
12. Subjective confidence in estimate	-243.00** (-97.14)	45.3 (-77.81)	-259.98** (-119.6)	-26.8 (-74.69)
13. Estimate		0.12** (-0.06)		
Constant	1317.07** (-569.07)	206.68 (-646.97)	956.86 (-853.56)	661.64 (-568.57)
N	130	130	130	130
R <sup>2</sup>	0.13	0.20	0.12	0.24

Note: Standard errors are reported in parentheses. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1%-level, respectively.

# Kvázi-experimenty

- Ako už bolo povedané, pri výskume vo verejných financiách nemôžeme pri každej otázke verejnej politiky použiť RCT. Diskutovali sme o alternatívach k RCT (časové rady, regresná analýza), ale zároveň sme uvideli, že tieto metódy majú nedostatky s odstraňovaním problému skreslenia.
- Je teda nejaká ďalšia možnosť ako vyšetriť kauzálne vplyvy aj bez použitia RCT?
- Empirický výskum vo verejných financiách sa v poslednom období začal zameriavať na riešenie, ktoré predstavuje akúsi strednú cestu: kvázi-experimenty. Ide o situácie, kedy zmeny v ekonomických podmienkach (napr. zmena zákona) prirodzene vytvoria takmer identické intervenčné a kontrolné skupiny. Pri kvázi-experimentoch je teda náhodné rozdelenie vytvorené namiesto experimentátora externými vplyvmi.
- Ani pri kvázi-experimentálnych štúdiách si nikdy nemôžeme byť istí, že z porovnania intervenčnej a kontrolnej skupiny sú vytlačené všetky skreslenia. Existujú však dva silné argumenty, prečo o kauzálnom efekte hovoriť môžeme. Prvý z nich je intuitívny - keďže intervenčná a kontrolná skupina boli pred zmenou veľmi podobné, je pravdepodobné, že skreslenia by mali byť minimálne. Druhý je štatistický - stále môžeme pre vyšetrenie skreslenia použiť alternatívne kontrolné skupiny.

# Kvázi-experimenty

- Metóda rozdielu v rozdieloch (difference-in-differences) — technika spájajúca časové rady a prierezovú analýzu, čím odstraňuje nedostatky oboch z nich. Porovnáваме zmenu v populácii A so zmenou v populácii B, pričom kontrolujeme faktory času, ktorý skresľuje analýzu časových radov v rámci populácie A. Tiež porovnáваме zmenu v rámci každej populácie, namiesto porovnania dvoch populácií v jednom bode, čím kontrolujeme skreslenia prierezovej analýzy dvoch populácií.
- Vyšetrenie zmeny premennej X
  - 2 obdobia (Y,Z)
  - 2 populácie (A,B)
  - v období Y je politika rovnaká pre A aj B
  - v období Z je nová politika pre A, zatiaľ čo politika pre B sa nemení
- $x(\text{populácia A, rok Y}) - x(\text{populácia A, rok Z}) = \text{Efekt intervencie} + \text{Skreslenie}$
- $x(\text{populácia B, rok Y}) - x(\text{populácia B, rok Z}) = \text{Skreslenie}$
- Rozdiel = Efekt intervencie

# Kváci-experimenty

■ TABLE 3-1

## Using Quasi-Experimental Variation

### Arkansas

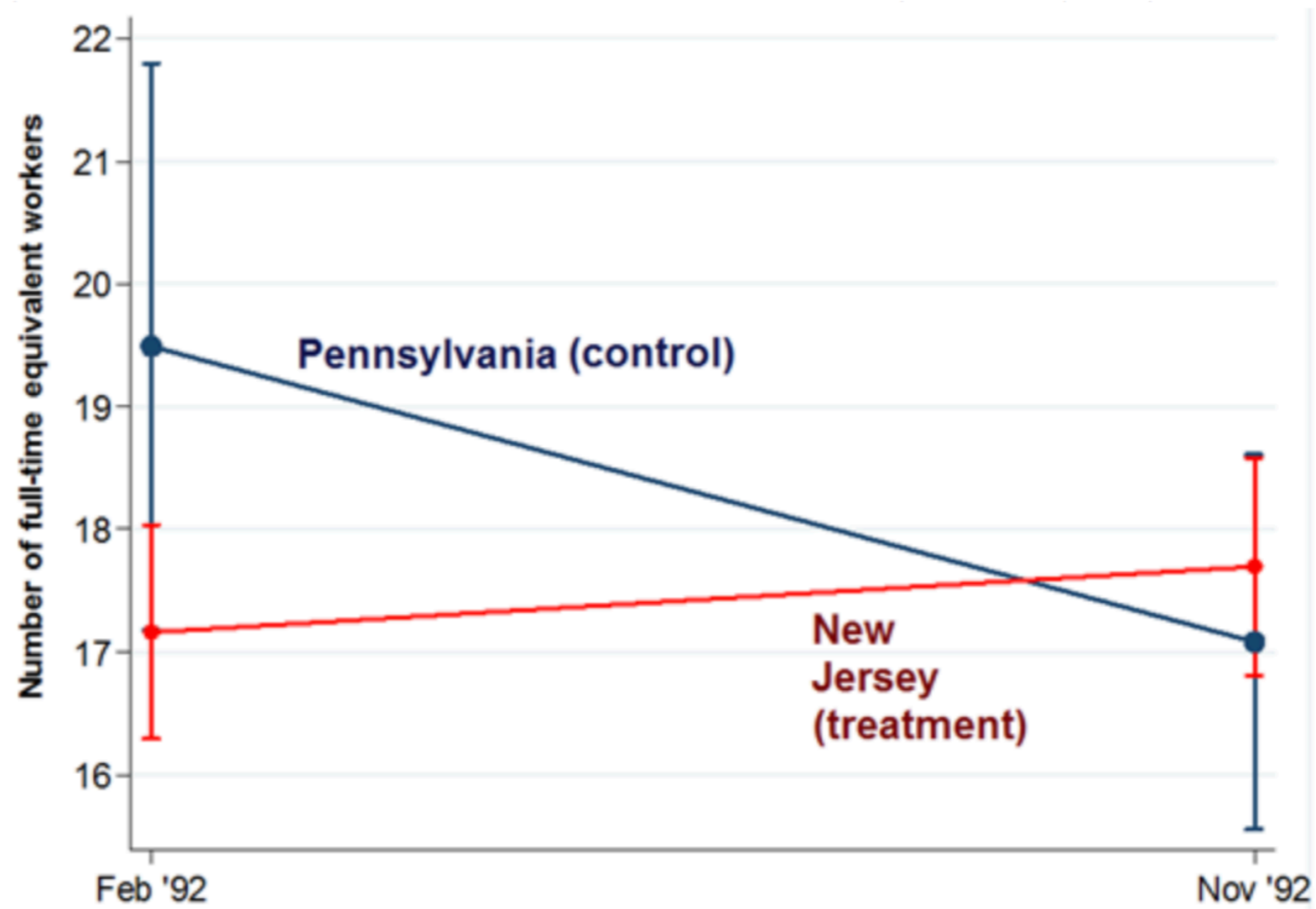
	1996	1998	Difference
<b>Benefit guarantee</b>	\$5,000	\$4,000	-\$1,000
<b>Hours of work per year</b>	1,000	1,200	200

### Louisiana

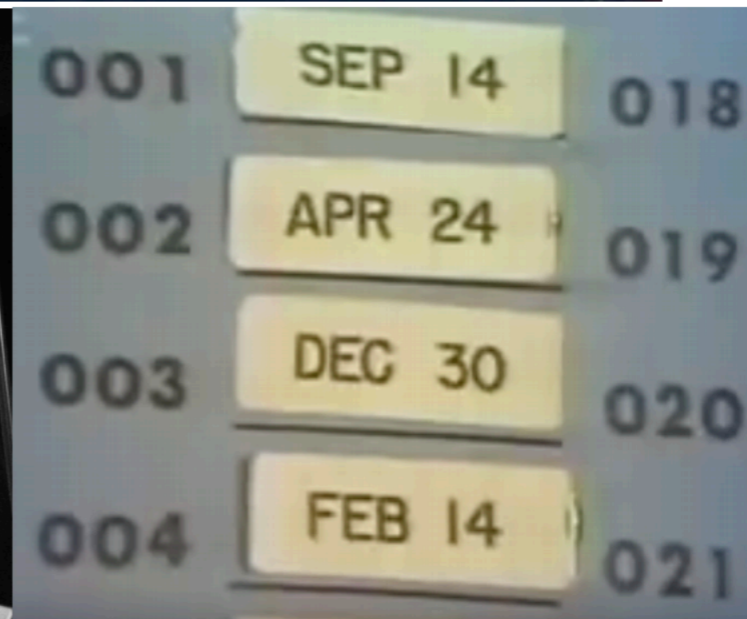
	1996	1998	Difference
<b>Benefit guarantee</b>	\$5,000	\$5,000	\$0
<b>Hours of work per year</b>	1,050	1,100	50

In Arkansas, there is a cut in the TANF guarantee between 1996 and 1998 and a corresponding rise in labor supply, so if everything is the same for single mothers in both years, this is a causal effect. If everything is not the same, we can perhaps use the experience of a neighboring state that did not decrease its benefits, Louisiana, to capture any bias to the estimates.

# Kvázi-experimenty - malo zvýšenie minimálnej mzdy v New Jersey negatívny dopad na zamestnanosť?

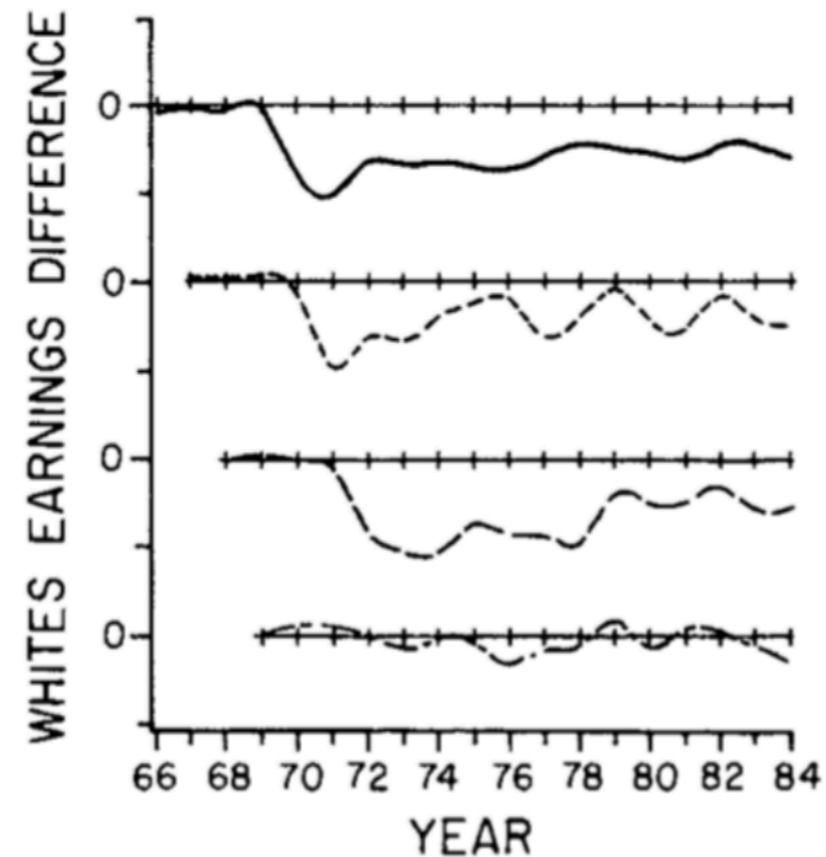
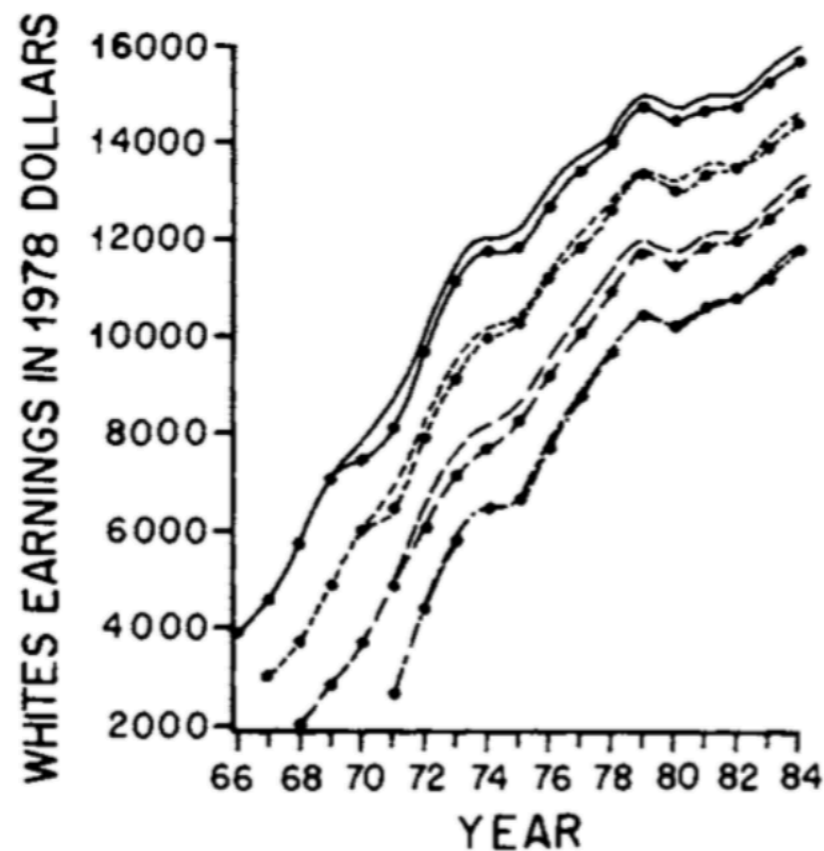


# Kvázi-experimenty - draft do vojny vo Vietname lotériou





# Kvázi-experimenty - draft do vojny vo Vietname lotériou



COHORT	DRAFT	
	ELIGIBLE	INELIGIBLE
1950	—●—●—●—	—————
1951	—●- -●- -●- -	- - - - -
1952	—●- -●- -●- -	- - - - -
1953	—●- -●- -●- -	- - - - -

COHORT	BORN
—————	1950
- - - - -	1951
- - - - -	1952
- - - - -	1953

# Obmedzenia experimentálnych pozorovaní

- Hoci RCT a kvázi-experimenty dokážu pomerne spoľahlivo odhaliť kauzálnosť, majú dve dôležité obmedzenia. Prvým je, že odhadujú iba kauzálny dopad konkrétnej intervencie. Povedzme, že sme zistili, že zníženie premennej A o 15 percent zvyšuje premennú B o 5 percent. To nám síce hovorí o efekte zníženia A o 15 percent, ale nehovorí to nič o tom, čo by sa stalo, ak by sme znížili A o napríklad 30 percent, alebo zvýšili o 10. Inými slovami, z pozorovania jednej zmeny v prostredí nevieme extrapolovať predpovede o iných zmenách v tom istom prostredí. Máme teda presnú odpoveď na jednu špecifickú otázku, ale nie všeobecný záver o tom, aké efekty by priniesli iné intenzity zmeny A.
- Druhým obmedzením je, že experimenty nám síce prezradia, čo intervencia zmení a ako, ale často nám nevedia povedať prečo - teda akým konkrétnym mechanizmom. Často by sme chceli poznať štrukturálne odhady, teda čo presne sa zmenilo v rozhodnutiach jednotlivcov (napríklad aby sme vedeli rozdeliť substitučný a dôchodkový efekt), ale experimenty nám dokážu poskytnúť iba tzv. redukované odhady.
- Redukované odhady ukazujú vplyv jednej zmeny na celkovú odozvu v správaní. Druhá nevýhoda experimentov je teda spojená s prvou - ak by sme dokázali rozumieť základnej štruktúre rozhodovania, potom by sme vedeli povedať viac o tom, ako budú ľudia reagovať na rôzne typy tej istej intervencie.

# Štruktúrálné modelovanie

- Problémy predostreté na predchádzajúcom slide viedli k vzniku štruktúrneho odhadovania. Touto metódou sa empirickí ekonómovia snažia odhadnúť nielen redukované odozvy na zmeny v prostredí, ale celé funkcie úžitku a to tak, že štiepia celkový efekt na viacero menších efektov.
- Štruktúrne modely sú potenciálne veľmi užitočným komplementom k experimentálnym analýzám, tieto modely sa však zostavujú oveľa problematickejšie ako redukované, pretože hoci používajú rovnaký objem dát, snažia sa z nich rozpoznať viac informácií.

# Kvantitatívna podpora kauzálnosti

- Model nemusí hovoriť pravdu, musí však byť užitočný
- Čím viac z nasledujúceho je splnené, tým spoľahlivejšie môžeme hovoriť o kauzálnosti
  - Efekt je výrazný. Nie, že je štatisticky významný ale že je veľký z praktického hľadiska.
  - Efekt je konzistentný. Podobný efekt namerali pre iné subjekty, v inom štáte za iných podmienok.
  - Efekt je špecifický. Skúmaný faktor ovplyvňuje len skúmanú odozvu  $y$  a nie 10 ďalších vecí.
  - Efekt rešpektuje časovú následnosť. Ak malo  $x$  spôsobiť  $y$ , tak  $x$  muselo nastať skôr.
  - Efekt je monotónny. Viacej  $x$  spôsobuje výraznejšiu zmenu v  $y$ .
  - Efekt je hodnoverný. Experti v danej oblasti poznajú mechanizmus pomocou ktorého mohlo  $x$  spôsobiť  $y$  a tento mechanizmus je hodnoverný.
  - Efekt je potvrdený experimentom.

# Sumár

- Základným cieľom empirickej práce je zdokumentovať kauzálne efekty jedného ekonomického faktoru na iný
- Problémom pri dosiahnutí tohto cieľa je, že vyžaduje intervenčnú skupinu zasiahnutú politikou a kontrolnú skupinu, ktorá zasiahnutá nie je. Zároveň musia byť (okrem politiky) tieto skupiny rovnaké. Ak rovnaké nie sú (systematicky sa líšia v niektorých charakteristikách), môže sa objaviť skreslenie - teda iné konzistentné rozdiely medzi skupinami, ktoré sú korelované s intervenciou, ale nie sú ňou spôsobené.
- RCT sú zlatým štandardom na riešenie problému skreslenia. Keďže sú pri nich už z definície intervenčná a kontrolná skupina rovnaké, analýza nie je skreslená a rozdiely v správaní vyvolané intervenciou sú kauzálnym efektom.
- Analýza časových radov zvyčajne nie je schopná poskytnúť presvedčivé odhady kauzálnych efektov, pretože v čase sa mení príliš veľa faktorov.
- Prierezová regresná analýza tiež trpí problémom skreslenia, pretože podobní ľudia sa môžu rozhodovať inak z dôvodov, ktoré nedokážeme pozorovať, čo tiež vedie k skresleniam. Zahrnutie kontrolných premenných ponúka potenciálne riešenie problému skreslenia.
- Kvázi-experimentálne metódy majú potenciál aproximovať RCT, ak sú kontrolné skupiny starostlivo vybrané.

- Prepokladajme, že Slovensko zaviedlo v roku 1999 zníženie dane. Chceme zistiť, či toto zníženie viedlo k nárastu osobnej spotreby. Máme nasledujúce údaje:

Year	Consumption in your state
1994	300
1996	310
1998	320
2000	350

- a. Spolužiak argumentuje, že najlepším odhadom efektu zníženia dane je zvýšenie spotreby o 30 jednotiek, ale vy si myslíte, že skutočný efekt je nižší, pretože spotreba stúpala aj pred znížením dane. Kto z vás bude mať lepší odhad?
- b. Predpokladajme, že v Česku sa daň počas sledovaného obdobia nezmenila. Zistili ste nasledovné:

Year	Consumption in neighboring state
1994	260
1996	270
1998	280
2000	300

- Aký bude váš odhad vplyvu zníženia dane na spotrebu teraz? Aké predpoklady potrebujete urobiť, aby ste mohli efekt odhadnúť správne?

# If you get this, you are ready for exam

