

**MIĘDZYNARODOWY ROK STATYSTYKI 2013**



**EKSPERYMENTY SONDAŻOWE:  
TYPOLOGIA I KRYTYCZNA ANALIZA  
ZASTOSOWAŃ W KONTEKŚCIE BADAŃ  
EWALUACYJNYCH**

**Włodzimierz Okrasa**  
*GUS & UKSW*

# Spis treści / Struktura prezentacji

## I. Wprowadzenie: *dlaczego eksperyment sondażowy?*

- ▣ dane eksperymentalne i dane obserwacyjne / sondażowe
- ▣ przyczynowość na poziomie zbierania i na poziomie analizowania danych
  - strategie rozumowań przyczynowych - ujęcia statystyczne vs. modele ekonometryczne

## II. Eksperyment sondażowy (ES) – pochodzenie i charakterystyka metody

## III. Zastosowania ES w badaniach metodologicznych i tematycznych (problemowych/*substantywnych*), w powiązaniu z:

- badaniami sondażowymi
- badaniami ewaluacyjnymi
- badaniami nad nierównościami i redystrubycją dochodów

## IV. Podsumowanie:

- ES w badaniach statystyki publicznej

# Tezy główne referatu:

3

## ▪ [Dlaczego eksperymenty sondażowe?]

**T:** Eksperyment Sondażowy stanowi proceduralny odpowiednik - na poziomie generowania danych - paradygmatycznych strategii rozumowań przyczynowych, gdy są one stosowane do *danych obserwacyjnych*, traktowanych na podobieństwo (faktycznie, zamiast) *danych eksperymentalnych*.

Założenie: to co wydarzy się w eksperymencie sondażowym odzwierciedla (i może być uogólnione na) to co wydarzy się w realnym świecie (społecznym)

→ TRAFNOŚĆ ZEWNĘTRZNA

# Dwie klasy strategii ujmowania przyczynowości w badaniach eksplanacyjnych i ewaluacyjnych w dyscyplinach nie-eksperymentalnych / dane obserwacyjne

4

- A . Strategie analityczne:
  - ▣ rozumowania przyczynowe na podstawie *danych obserwacyjnych* w ramach *kontrfaktycznych* schematów relacji przyczynowo-skutkowych;
  
- B. Strategie empiryczne:
  - ▣ generowania danych o charakterze eksperymentalnym w badaniach (dyscyplinach) nie-eksperymentalnych / w bad. społecznych, w tym w bad. ewaluacyjnych dot. ‘przyczyn efektów’ oraz/lub ‘efektów przyczyn’.

# Ad. Strategie analityczne

5

Dwa typy strategii ‘rozumowań przyczynowych’ w literaturze::

- (i) Statystyczne: odwołujący się do pojęcia ‘wyniku potencjalnego’ ***kontrfaktyczny paradygmat Neymana-Rubina*** (*Rubin, 1974, Holland 1986; Rosenbaum 2002; Okrasa 2012*)
  - o brak specyfikacji modelu wpływu & bezwarunkowe założenia dot. losowości przyporządkowań jednostek (badanych) do źródeł zmienności (‘interwencji’) i do możliwych efektów (stanów kontrfaktycznych)

[„modelowanie efektów przyczyn, a nie przyczyn efektów” *Holland 1986*]

- (i) Ekonometryczne: zorientowany na artykulację mechanizmu (w tym manipulacji i wyboru stanu kontrfaktycznego) generowania ‘efektu’ */treatment effect*) ***ekonometryczny model przyczynowy*** rozwijany w ramach ewaluacji wpływu, uwzględniający błąd (samo)selekcji w ocenie wpływu ‘interwencji’ (*Heckman 2001; Heckman i Vytlacil 2007*).

# Kontrfaktyczny schemat przyczynowości w statystyce (N-R)

6

- Przyjmując oznaczenia:  $Y_i^t$  oraz  $Y_i^c$  jako wyniki zmiennej losowej dla jednostek w grupie poddanej działaniu (eksperymentalnej) i niepoddanej (kontrolnej), odpowiednio, , jednostkowy efekt oddziaływania wynosi:

$$\delta_i = Y_i^t - Y_i^c .$$

- Obserwowany wynik zmiennej:

$$Y = Y^t \text{ if } D = 1$$

$$Y = Y^c \text{ if } D = 0$$

albo:

Grupa 'eksp.' /poddana działaniu ( $D = 1$ ) &  $Y = Y^t \rightarrow$  Obserwowane jako  $Y$

Grupa 'eks.'/niepoddana działaniu ( $D = 0$ ) &  $Y = Y^c \rightarrow$  Warunek **kontrfaktyczny**

Grupa kontrolna ( $D = 0$ ) &  $Y = Y^t \rightarrow$  Warunek **kontrfaktyczny**

Grupa kontrolna ( $D = 0$ ) &  $Y = Y^c \rightarrow$  Obserwowane jako  $Y$

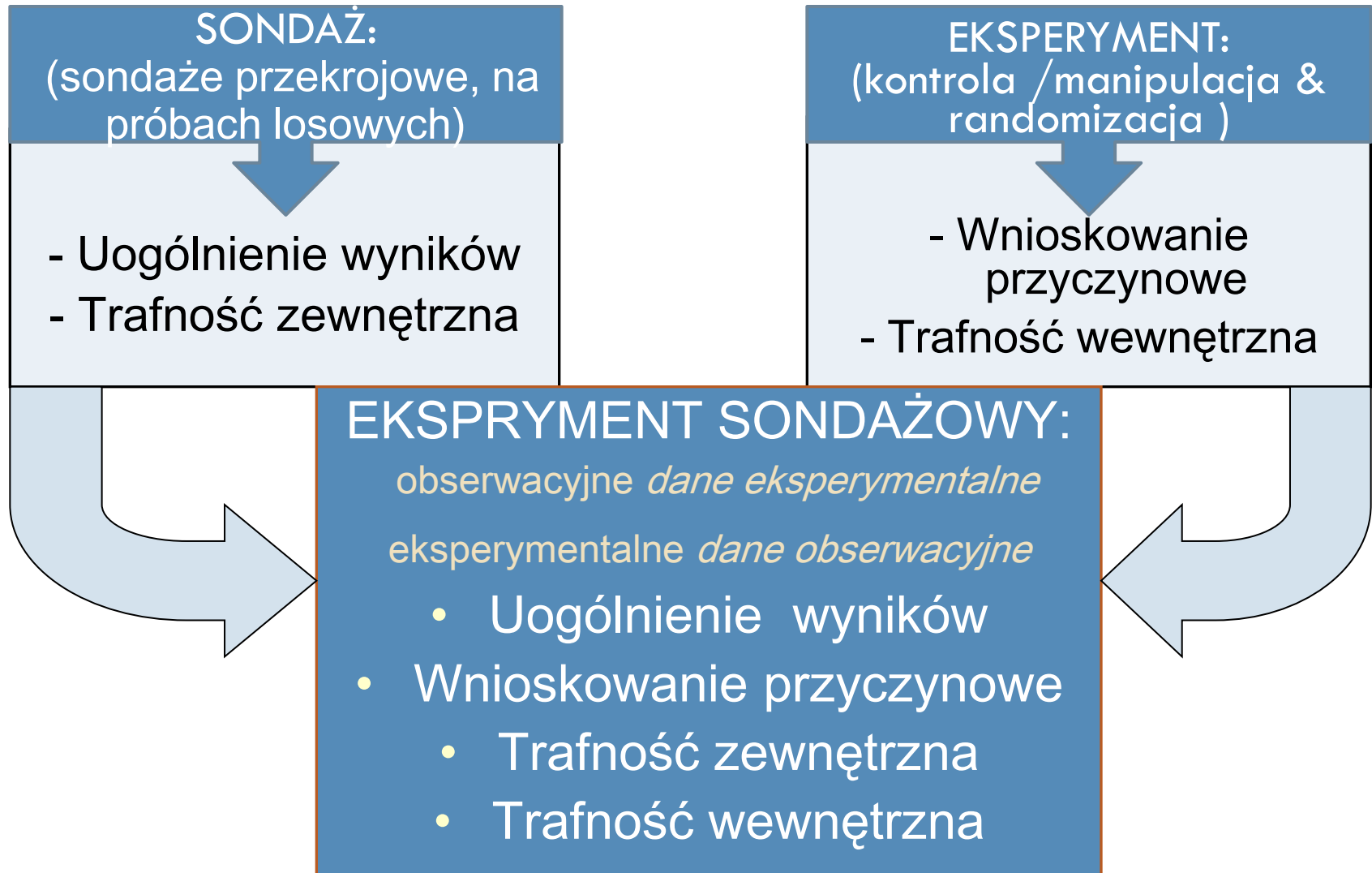
Zagregowany efekt działania ('interwencji') w populacji / ATE:

$$E(\delta) = E[Y^1 - Y^0] = E[Y^1] - E[Y^0]$$

$\rightarrow$  ES jako sposób uwzględnienia / konstruowania warunku kontrfaktycznego (w analizach przyczynowych – obok eksperymentów klasycznych i innych (terenowych, naturalnych).

SE jest jakościowo inna/'nową' metodą  
– nie będącą zwykłą hybrydą *sondażu* i *eksperymentu* –  
łącząca własności obydwóch

7



# ES: Charakterystyka metody

8

- Eksperyment ('klasyczny' / laboratoryjny): warunki realizacji badania:
  - wprowadzenie oddziaływania (*treatment*) interwencji, bodźca (*stimulus*), doświadczenia;
  - losowe przyporządkowanie badanych do bodźca /grupy eksperymentalnej i do grupy kontrolnej
  - pomiar wyniku
    - pomiar zmiennej 'wynikowej' przed- i po- zastosowaniu bodźca
    - różnice wyników w gr. eksperymentalnej i kontrolnej traktowane jako efekt 'doświadczenia' (bodźca);
  - brak zakłóceń trafności wewnętrznej.



# ES: charakterystyka metody – kont.

9

- Wprowadzenie systematycznego źródła zmienności - jako elementu sondażu – pomiędzy badanymi
- Brak (zazwyczaj) pomiaru przed- ‘bodźcowaniem’ (*pre-testu*) / brak porównań ‘podłużnych’;
  - odmienność SE od eksperymentu klasycznego
    - pomiar po-/(*ex post*) tylko -- grupa kontrolna: porównania ‘równoległe’
- Losowe przyporządkowanie badanych do grup (do ‘bdźca’/interwencji i do gr. kontrolnej)
  - [bodźcowanie – najczęściej w formie przekazu informacji/instrukcji, drogą kontaktu mieszanego: internetowo-komputerowo/mailowo – telefonicznego]

# ... charakterystyka metody – kont.

10

- zakłada się możliwość wnioskowanie przyczynowego (spełniającego 3 p-wowe wymagania dot. relacji przyczynowej) :
  - (1) obydwie zmienne współwystępują w czasie lub w przestrzeni,
  - (2) zm. przyczynowa musi poprzedzać zm. wynikową/efekt w czasie,
  - (3) relacja pomiędzy ‘przyczyną’ i ‘skutkiem’ nie może być objaśniona działaniem jakiejś trzeciej zmiennej (jak w korelacji pozornej)
- eksperymenty sondażowe nie tylko eliminują problem trzeciej zmiennej, ale wykazują przewagę zarówno nad coraz powszechniej stosowanymi sondażami internetowymi, z jednej strony, oraz nad eksperymentami laboratoryjnymi w naukach społ. (typu RCT) z drugiej strony.

# Zalety stosowania ES w porównaniu do tradycyjnego eksperymentu ('laboratoryjnego') w bad. społ.

11

- (i) **próbka z populacji zamiast grupy (podobnych sobie) osób** (zazwyczaj studentów) - możliwość przyporządkowania losowo wybranych podmiotów (jedn. bad.) do odpowiednich bodźców, również w sposób losowy używając technik próbkowania, na odpowiednio dużą skalę i w odpowiednio zróżnicowany sposób;
- (ii) **bliżej realnego świata** niż warunki laboratoryjne – szansa na większą trafność zewnętrzną;
- (iii) **koncentracja** badania na określonej kategorii osób w populacji, szczególnie w badaniach panelowych z użyciem technik internetowych;
- (iv) **zwiększenie znaczenia badań typu eksperymentalnego** w n. społ., co prowadzi do wzrostu poziomu metodologicznego nauk społ., m. in. dzięki możliwościom porównawczej ewaluacji alternatywnych hipotez 'przyczynowych'; także jako etap następny po badaniach sondażowych (np., czy wzrost zaufania społecznego prowadzi do wzrostu zakupów poprzez internet czy vice-versa - badania potwierdzają zależność dwukierunkową)

# Typologia zastosowań eksperymentów sondażowych<sup>\*)</sup>

12

Cele badań	<i>Badania podstawowe</i>	<i>Badania stosowane &amp; ewaluacyjne</i>
<b>Metodologiczne</b>	Testowanie mocy wyjaśniającej teorii - np. t. <i>postmaterializmu</i> Ingleharta w krajach o wys. bezrobociu i o wys. Inflacji – rozbieżne wyniki (Clark i in., 1999)	Alternatywne podejścia dla : - zwiększenia stopy odpowiedzi - zmniejszenia obciążenia resp., - usprawnienie procesu ankietowania /top-to-bottom - polepszenie jakości danych
<b>Substantywne</b>	Testowanie zasięgu teorii : np. t. wzrostu uprzedzeń etnicznych – pyt. o akcję afirmatywną i o postawę do kolorowych	Testowanie wyniku /porównania Porównawcza ewaluacja programów /wyznaczanie złożone (czynnikiowe)

\*) Źródło: Na podstawie S.L. Nock i T. M. Guterbock (2010).

# Ad. Strategie empiryczne: geneza eksperymentów sondażowych (≠ sondaże eksperymentalne)

13

- **Historycznie & dyscyplinarnie:**
- **n. polityczne** - Paul Sniderman i in. (*Survey Research Center at the University of California at Berkeley*) z wykorzystaniem technik CATI i *onlin'*owych, CAWI, CASI, pochodnych i ich kombinacji;
  - ▣ bad. opinii publ. - dlaczego ludzie nie lubią mediów?; b. nad terroryzmem, itp.
- **n. ekonomiczne** – badania porównawcze nad rynkiem pracy/satysfakcją z pracy & subiektywnym dobrostanem; nad nierównościami i stos. do redystrybucji doch.,
  - ▣ satysfakcja z pracy jako element dobrostanu (Clark 2006; OECD 2013 – także w kontekście reform syst. emerytalnych / skłonności do pracy dłużej w UE);
- **n. społeczne** – badania interdyscyplinarne i transdyscyplinarne:
  - ▣ bad. nad sondażami / nad metodami bad. sond. → ES główną strategią
  - ▣ bad. ewaluacyjne & bad. nad ewaluacją / metodologia b. ewaluacyjnych;
- **n. behavioralne:** porównania interpersonalne w perspektywie między-kulturowej (pojęć i wartości, etc.) - bad. nad ryzykiem, postrzeganiem zagrożeń, itp.

# Rodzaje danych i analiz

14

	A. Analiza sondażowa	B. Analiza eksperymentalna / przyczynowa
1. dane obserwacyjne	✓	
2. dane eksperymentalne		✓
3. 'eksperymentalne' <i>dane obserwacyjne</i>	✓	✓
4. 'obserwacyjne ' <i>dane eksperymentalne</i>	✓	✓

## ... rodzaje danych i analiz – *kont.*

15

a) *dane obserwacyjne*

→ analiza sondażowa

b) *dane eksperymentalne*

→ analiza przyczynowo-skutkowa

c) *'eksperymentalne' dane obserwacyjne:*

(i) analiza sondażowa: SONDAŻ z eksperymentem 'zanurzonym'

- usprawnienie narzędzi -- porządek i sformułowanie pytań
- porównywalność międzykulturowa (np. pojęć w ESS: 'klasa średnia' - 'wyższa' – 'niższa' (Hopkins & King, 2010)

## ... rodzaje danych i analiz – kont.

16

→ **techniki winietowe** /techniki winiet zakotwiczonych (*anchoring vignettes* = dodatkowe itemy z hipotetycznymi obiektami / sytuacjami (Bruch & Mare 2012)

-- 'efekt kontekstu' sondażu (Buckley 2008)

■ kierunkowy vs. korelacyjny efekt kontekstu sondażu (Tourangeau i in. 2000)

■ efekty bezwarunkowe i warunkowe (Smith, 1986)

- czy nowe itemy / pytania będą 'pracować'?

(ii) analiza eksperymentalna:

- określenie kierunku i/lub wielkości 'efektu doświadczenia'

(*treatment eff.*) -- np. w procesach formowania opinii i postaw

- akcja afirmacyjna i stereotypy: kolejność pytań w dwóch grupach a negatywne postawy;

→ relacja przyczynowo-skutkowa jeśli w sondażu to także w *realu*).



# ... rodzaje danych i analiz – kont.

## d) 'obserwacyjne' dane eksperymentalne

### (i) analiza sondażowa:

- porównanie wzorów & wyznaczników preferencji i wyborów: np. w zakresie miejsca zamieszkania (Bruch & Mare 2012); na bazie winiet przedstawiających kombinację 14 elementów otoczenia rasowego d. miejsca – od „wszystkie domy dookoła zamieszkałe przez etnicznie ‘innych’ (w USA ‘czarnych’) do odwrotnie (przez wszystkie domy ‘białe’ – 5 kombinacji/wzorów = pokazywanych białym i kolorowym)

### (ii) analiza eksperymentalna:

- weryfikacja hipotez nt. zależności przyczynowo-skutkowych --  
oszacowanie 'efektu bodźcowania' w 'populacji' pot. beneficjentów  
→ badania ewaluacyjne & badania nad ewaluacją  
(metodologia ewaluacji)

# ES odpowiedzią na wyzwania wobec prób wnioskowania przyczynowego z danych sondażowych (przekrojowych)

18

## □ Wyzwania:

- obciążenie selekcją
- korelacja pozorna
- korelacja błędów pomiarowych
- obcięcie danych w czasie (*censored data*)
- brak ‘prawdziwych’ kontrfaktów
- wzajemne współzależności przyczynowe

[Np., powszechne rozbieżności w porównaniach ocen polityk/programów rządowych na podstawie danych sondażowych oraz danych eksperymentów naturalnych – użycie technik łączeniowych/*matching*, *propensity score* czy dwu-stopniowej regresji może zredukować obciążenie, ale go nie eliminuje (Gaines, Kuklinski i Quirk, 2006→(GKQ))].

Dane panelowe trudno dostępne i także niepozbawione błędów ze wzgl. na ubytki (*attrition*) oraz możliwą korelację błędów w czasie i efekt ‘uczenia się’ (respondenta).

# ES w metodologii badań sondażowych

19

## □ Usprawnienia w zakresie:

- obniżenie kosztów – szybkie dostarczenie wyników – zwiększenie użyteczności badania – zmniejszenie błędu sondażowego – procesu wywiadu & jakości danych, m. in., niedoszacowania dochodów & uczestniczenia w programach socjalnych
  - w szczególności, redukcja:
    - błędu próbkowego & błędu pokrycia (porównania np. losowanie adresów vs. jedn. terytorialnych; albo RDD vs. lista mieszkań, etc.;
    - błędu związanego z brakami danych (np. Dillmana TDM) i z brakami odpowiedzi (np. wizualizacja, systemy pominięć, inne ułatwienia)
    - błędu pomiarowego związanego z narzędziem (stopień ustrukturywania, sformułowania i sekwencja pytań, etykiety w skalach, itp.) i z ankieterem (sposoby raportowania, ‘wywiady poznawcze’, itp.)
- Efektywność ES większa w ramach sondażu panelowego (np. SIPP) & w eksperymentach złożonych (‘czynnikiowych’)

# ...ES w metodologii bad. sond. – kont.

20

- Typ eksperymentu w ramach ES wg schematu wyznaczania:
  - w. proste – porównanie np. kilku różnych wersji kwestionariusza (efekt konstrukcji) lub programu (zmiany w pr.)
  - w. złożone / czynnikowy – efekt wielu zmiennych w jednym badaniu, wielo-doświadczeniowym {np. układ pytań ⊗ treść pyt. ⊗ sposób administrowania} oraz efekt *interakcji* pom. czynnikami
  - ryzyko ‘eksplozji kombinatorycznej’ (dla 3 zm. o 3 poz. = 27 możliwości, stąd wyznaczanie *frakcyjne* (tylko część spośród  $2^k$  komórek, wystarczy żeby była zachowana proporcjonalność
    - np. w badaniach (Nock’a 2010) z 10 czynnikami od do 2 9 kategorii każdy otrzymano 270,000 możliwych kombinacji winietowych.]

## ...przykłady zastosowań (w bad. metod.) – kont.

21

- **Zwiększenie stopy odpowiedzi** poprzez eksperymentalne dostosowanie sondażu do skłonności potencjalnych respondentów do uczestniczenia w badaniu:
  - ▣ sondaże postaw konsumenckich: eksperyment z losowo dobranymi g.d. (3000 gd), badanymi sekwencyjną kombinacją technik w ramach komputerowo-internetowo-pocztowo-telefonicznych strategii mieszanych – z intencją dostosowania sondażu do grup respondentów (*tailored survey design*) mniej i bardziej skłonnych do uczestniczenia w badaniu, bazując na *paradanych* (informacjach z wcześniejszych sondaży nt. przyczyn braku danych /br. odpowiedzi (Luiten & Schouten, 2013)
    - ➔ znaczący wzrost uczestnictwa i reprezentatywności (przy nieznacznym wzroście kosztów z racji preferowania pocztowego kontaktu nad internetowy, ale mniej kosztownego niż CATI, o który należało uzupełnić wersję pocztową dla wzrostu odpowiedzi)

## ...przykłady zastosowań (bad. metod.) – kont.

22

- **Redukcja braków danych w badaniach longitudinalnych /s. panelowych (*sample attrition: British Household Panel Survey*):**
  - 1wszy experiment, porównujący strategie dot. sposobu kontaktowania się i informowania badanych – skłonność do uczestniczenia w kolejnej rundzie badania okazała się wzrastać dzięki między-rundowej wysyłce karty pocztowej z prośbą o informację nt. ew. zmiany adresu (karta o potwierdzenie adresu tak jak brak wysyłki karty, nie podnosił udziału);
  - 2gi eksperyment porównywał standaryzowane i dostosowane do respondentów formularze dla ludzi młodych oraz dla osób b. zajętych – efekt pozytywny nieznaczny w obu grupach, natomiast opcja w postaci krótkiego wywiadu telefonicznego znacząco polepsza w stosunku do f2f wywiadu (Fumagalli i in., 2013).

# Problemy: kontaminacja wielo-doświadczeniowego badania eksperymentalnego w ramach tego samego sondażu populacji (GKQ):

23

- dla wyjaśnienia dlaczego program jest/nie jest efektywny (ma/nie ma efektu w wymiarze społ.politycznym) ES powinien pozwolić na ocenę długo-czasowego efektu
  - ▣ trudność w przypadku wielo-doświadczeniowego eksperymentu (eksperymentu złożonego), w szczeg. realizowanego sekwencyjnie: doświadczenia wcześniejsze wpływają na późniejsze eksperymenty - *spillover* efekt, tym większy im więcej doświadczeń w ramach eksperymentu
- w sondażu wielo-itemowym z losowym przyporządkowaniem podmiotów do grup kontrolnej i eksperymentalnych (bodźcowanych), dwa razy podział uczestników eksperymentu: (i) pytanie o to czy kolorowi są leniwi w doświadczeniu 1-wszym, w którym badany ma 50-procentową szansę otrzymać przed pytaniem informację nt. akcji afirmatywnej (AA); (ii) paralelnie (doświadczenie 2gie): pytanie o brak odpowiedzialności kolorowych z taką samą szansą otrzymania informacji nt. urodzeń poza-mażeńskich w tej grupie; grupa kontrolna nie otrzymuje informacji dodatkowych (bodźców).

# Problemy – spillover efekt – kont.

24

- przyjmując:  $b$  – pr-wo bazowe podzielenia stereotypu, zaś  $p$  i  $q$  dodatkowe wielkości, o które pr-wo bazowe może być zwiększone w wyniku doświadczenia 1-go i 2-go, odpowiednio; w odróżnieniu od porównania tylko różnic pom. wielkościami w grupach eksperymentalnych i kontrolnej można rozpatrywać łączne pr-wo dla oszacowania efektu *spillover*, przyjmując,  $y_k$  jako proporcję dzielących stereotyp  $k$ :

□

		T2		
		0	1	
T1	0	$n_{00}$	$n_{01}$	$n_{0.}$
	1	$n_{10}$	$n_{11}$	$n_{1.}$
		$n_{.0}$	$n_{.1}$	$N$



# Problemy – *spillover* efekt – kont

25

- Oczekiwany efekt bodźca w 1 wszym doświadczeniu:  
 $E[(y_1|T1=1) - (y_1|T1 = 0)] = n_{11}(b+p)/n_{1.} - n_{01}(b)/n_{0.} = p.$
- oraz dla doświadczenia 2-go

$$E[(y_2|T2=1) - (y_2|T2 = 0)] = [n_{11}(b+p+q) + n_{01}(b+q)]/n_{1.} - [n_{10}(+p) + n_{00}b]/n_{0.} = q + [n_{11}/n_{1.} - n_{10}/n_{0.}] p$$

efekt możliwego oddziaływania doświad. 1-go na wynik zarejestrowany w dośw. 2gim (tym większy im większe odstępstwo od losowości w wyznaczaniu uczestników do dośw. 2-go); formalnie od  $n_{i1}$  i od  $n_{k0}$  (liczby doświadczeń z

dwójkowym przyporządkowaniem, czyli  $\begin{matrix} k \\ 2 \end{matrix}$

W badaniach ewaluacyjnych, czy nad innowacyjnością, analogiczny problem do *spillover* jest realny (np. dyfuzja innowacji przestrzenna).

# 'Tematyczne'(problemowe) przykłady zastosowań

26

Populacyjne Eksperymenty Sondażowe (*PES*) / *Population-based Survey Experiments*) D. Mutz (2011) – eksperymenty ‘zanurzone’ w sondażu (ogólno-krajowym) typu online (Druckman i in. 2006).

- Eksperymentalne Badania nad Dzielaniem się Czasem w Naukach Społecznych / *Time-Sharing Experiments for the Social Sciences (TESS)*, w ramach NSF Projektu Wielu-Badaczy (Multi-Investigator Project) -- promowanie metodologicznej innowacji w postaci PES:
  - małe grupy badaczy realizujące telefoniczny sondaż, zaprojektowany jako ‘eksperyment’ - wykorzystując jako wspólne: (i) moduł podstawowy zawierający pytania socjo-demograficzne ale (ii) odmienne, własne moduły tematyczne [forma integracji sondażowej na poziomie zbierania danych] – projekt realizowany od pocz. 2000 umożliwiający realizację PES na skale globalną; ExperimentCentral.org]

# Zastosowania w kontekście nierówności i redystrybucji: Analiza zależności poparcia programów redystrybucji od postrzeganej nierówności rozkładów dochodów (NBER 2013)

27

- Efekt informacji nt. nierówności i podatków na stosunek do redystrybucji stosując SE (onlinowy serwis *Amazon Mechanical Turk /mTurk* \$1-2 per respond.): 5,000 respondentów losowo przyporządkowanych do ‘bodźców’ polegających na podaniu informacji nt. nierówności dochodów w USA oraz związku pom. opodatkowaniem wysokich dochodów a wzrostem gosp.,
  - → b. znaczący wpływ informacji na postrzeganie nierówności przez badanych ale bez konsekwencji w postaci poparcia dla programów redystrybucji (ani podatkowej ani świadczeniowej)
  - Wyznaczanie ES: (1) 6 rund panelowych (w l. 2011-2012) z ‘core vars’ (socio-ekonomiczne i zatrudnienie); (2) informacja nt. nierówności i podatków w gr. ‘doświadczalnych’; (3) pytania o postrzeganie nierówności, politykę podatkową i transfer dochodów. Dane z mTurk porównano (*matching*) z danymi z CBS-News) na 32 ‘czynnikach’: płeć (2) x wiek (2) x rasa (2) x wykształc./college (2) x głosowanie na prezydenta /Demokratę (2)

# ES w badaniach ewaluacyjnych

## - efektywność programu a obciążenia postaw wobec polityki redystrybucyjnej :

28

- Obciążenie ocen i poparcia (braku poparcia) wobec programów redystrybucji dochodów w zależności od ‘grupy odniesienia’ z jaką identyfikuje się respondent:
  - ▣ eksperyment w ramach sondażu z informacją o pozycji badanego na rangowej skali dochodu
  - losowo wybrani (w ramach większego sondażu) oceniając rozmiar redystrybucji do odpowiednich grup zmieniali swoje opinie po skonfrontowaniu (w kolejnym doświadczeniu) swoich ocen nt. własnej pozycji w rankingu dochodów z otrzymaną w badaniu ‘pozycją ‘obiektywną’ – im większa rozbieżność w dół (od nad-szacowanej pozycji własnej/poz. ‘bogactwa’) tym większe poparcie dla wzrostu (zakresu i poziomu) redystrybucji [nb., potwierdzenie t. Rawlsa sprawiedliwego podziału /zasady *maximinu*.]

# Badania nad ewaluacją:

- studia nad praktykami i procedurami ewaluacyjnymi

29

- Wyznaczanie eksperymentalne jako rygorystyczne podejście do oceny efektywności interwencji z losowym przyporządkowaniem badanych do grup doświadczalnych i kontrolnych;
  - ewaluacja na bazie wyznaczania (quasi)eksperymentalnego powinna pozwalać na ocenę oszacowania wpływu ‘czystego’ (uwzględniając potencjalne zagrożenia) – stąd oczekiwania (sponsorów interwencji) dot. stosowania na możliwie szeroką skalę schematu losowej kontroli prób /RCT (Random Control Trial) – z rezerwą do RCT (popularnego w ewaluacyjnych bad. kinicznych) jako ‘złotego środka’ podchodzi EES /European Evaluation Society (nie rekomendujące żadnej konkretnej procedury w tej roli);
    - dyskusja trwa nad kwestią czy grupa kontrolna wystarcza dla skonstruowania kontrfaktycznej wersji interwencji (jak użycie RCT)

# ... ES w kontekście ewaluacji

30

- - wyznaczanie na bazie losowości i kontroli (jak RCT) nie może wg ekspertów ewaluacji być uznane za najlepsze podejście pozwalające na analizy przyczynowo-skutkowe - np. Scriven uznaje quasi eksperyment ('naturalny') za metodologicznie równoważny RCT, ale
  - RCT za praktycznie nieprzydatny, nietrafny dla eliminacji korelacji pozornej w eksplanacji bo trudność w spełnieniu zasady podwójnego zaślepienia (*double-blind*), gdy tymczasem powinno być 'pełne-zaślepienie' dla oddzielenia efektu 'interwencji' od psychologicznego efektu otrzymywania 'bodźca' (interwencji) w ewarunkowaniu przyczynowym; nawet najpoprawniejsze badanie wg RCT powinno być uzupełnione o 'test polowy; '
  - RCT bywa ceniony za jego rolę (przydatność) w kontrfaktycznych rozumowaniach przyczynowych (w projektowaniu stanów kontrfaktycznych)
    - w praktyce badawczej ES preferowane (bliższe rzeczywistości)

# PODSUMOWANIE:

31

- ES w badaniach statystyki publicznej
  - ▣ badania metodologiczne / nad sondażem – obszary rekomendowane do badań za pomocą ES:
    - problem obciążenie respondentów
      - redukcja poprzez sondaże ‘dopasowane’ do kategorii resp. (‘tailored survey’);
    - badania nad podniesieniem jakości
      - lepsze wyznaczanie próby w kolejnych rundach badań stałych
      - ulepszenie narzędzi i sposobów zbierania danych / dostosowane do kategorii respondentó (bazując na wykorzystaniu paradanych /sondaże ‘dopasowane do respondenta’))

# ...podsumowanie – kont.

32

- ▣ badania ewaluacyjne / bad. nad ewaluacją – niewykorzystane możliwości usprawnienia i nowe obszary badawcze (w obszarze bad. ewaluacyjnych):
  - bad. nad wyborami w zakresie (poparcia dla) programów lokalnych włączających ‘uczestników’ programów/*stakeholders* (identyfikacja wyznaczników preferencji i sposobów ‘perswazji’ w procesach decyzyjnych – w szczególności nt. dóbr publicznych i obciążeń postaw ‘uczestników’;
    - -- także w badaniach ewaluacyjnych *ex ante* nad innowacyjnymi przedsięwzięciami lokalnymi (efektywnością bodźców inicjujących takie przedsięwzięcia).



*Dziękuję za uwagę*  
*!*