

Planowanie eksperymentów

Michał Krawczyk

Wydział Nauk Ekonomicznych UW

Kiedy reszta świata uzna, że nasze pomysły są interesujące?

- Projekty motywowane przez literaturę teoretyczną
 - Przykład: Eeckhoudt (1996) – „powściągliwość” w podejmowaniu ryzyka
- Projekty motywowane przez literaturę eksperymentalną
 - (np. niejednoznaczność wyników poprzednich badań eksperymentalnych danego zagadnienia)
 - Przykład: Fischbacher i in. (2001) – altruizm vs. warunkowa kooperacja w grze w dobra publiczne.
- Projekty wyjaśniające niebanalne zjawiska obserwowane w terenie („*stylized facts*”).
- Przykład: dlaczego płace przekraczają WTA pracowników? „wymiana darów”

Przykład formułowania pytań badawczych: Krawczyk (2010)

- Amerykanie akceptują znaczne nierówności społeczne
- Europejczycy opowiadają się za redystrybucją
- Artykuły Fong, Alesiny i innych: postrzegana równość szans może sprzyjać akceptacji nierówności wyników (jest to jeden z aspektów tzw. sprawiedliwości proceduralnej)
- Modele teoretyczne podejmowania decyzji społecznych w warunkach ryzyka przewidujące preferencje dla równości *oczekiwanych*, nie (tylko) faktycznych wypłat (Trautmann 2009, Krawczyk 2011).
- Niewiele wcześniejszych badań eksperymentalnych
→Projekt miał szanse

Jak szukać literatury

- **Kolejność:**
 - **Książki**
 - **Artykuły przeglądowe, metaanalizy (np. w JEL)**
 - **Artykuły raportujące poszczególne badania**
- **Metoda kuli śniegowej**
(wg cytowań, wg autorów)

Przygotowywanie planu eksperymentu

- **Czy metoda eksperymentalna najwłaściwsza?**
- **Przykład negatywny: zachowania spadkowe**
- **(choć ekonomia eksperymentalna bada coraz dziwniejsze rzeczy, m.in. jak alkohol wpływa na decyzje ekonomiczne a podniecenie na gotowość podejmowania ryzykownych kontaktów seksualnych)**

Przygotowywanie nowego designu

- Główne pytanie: czego dokładnie chcesz się dowiedzieć z przeprowadzonego eksperymentu?
- Zły pomysł: maksymalnie upodobnić środowisko podejmowania decyzji tworzone w labie do sytuacji społecznej, która jest przedmiotem naszego zainteresowania i „zobaczyć co się stanie”.
- Celem eksperymentu nie jest udawanie rzeczywistości zewnętrznej
- Jest nim testowanie hipotez wynikających z teorii
- Przykład: eksperymentalna giełda może być względnie prosta jeśli teorii milczy nt. tego jak na zachowanie wpływa złożoność sytuacji

Określenie bazowej decyzji badanego

- Typowa pojedyncza decyzja badanego winna zależeć od preferencji i okoliczności, o których teoretyzujemy a nie zależeć od innych
- (Np. jeśli teoria dotyczy zdolności poznawczych, decyzje nie powinny zależeć od preferencji społecznych, stosunku do ryzyka itd.)
- Ważne rozstrzygnięcie: czy (jak) decyzje badanego będą miały wpływ na innych
- (Czasem odpowiednio dobrane decyzje indywidualne jako warunek kontrolny dla interakcji, np. decyzja w warunkach ryzyka jako analogon decyzji społecznej o powierzeniu innemu graczowi pieniędzy w „grze w zaufanie”)

Przykład planowania designu: Krawczyk (2010)

- Idea: „równość szans” wpływa na „preferencje dla redystrybucji”.
- Założenia do designu:
 - Niepewność dochodów
 - Nierówności dochodowe
 - Większa lub mniejsza równość szans osiągnięcia wysokich dochodów
 - Istnieje redystrybucja
- Ważny punkt: odróżnienie skłonności do redystrybucji od interesu własnego i stosunku do ryzyka
- (nierówność szans może oznaczać, że silniejsza redystrybucja bardziej się opłaca, podczas gdy interesująca nas teoria skupia się na innym mechanizmie. Istotnie, ktoś może wspierać redystrybucję właśnie dlatego, że jego indywidualne szanse osiągnięcia sukcesu są niewielkie.)

Rozwiązanie: plan badawczy

- Badani napotykali zróżnicowane indywidualne prawdopodobieństwa „sukcesu”, polegającego na otrzymaniu raczej 30 euro niż nic.
- 4-osobowe grupy w każdej z rund
- W każdej grupie dokładnie dwie osoby odniosą „sukces”.
- W każdej z rund badani głosowali za preferowanym poziomem redystrybucji (od tych, którzy zdobędą 30 euro do pozostałych)
- Ponieważ skład grup się zmieniał, więc wewnątrzgrupowa wariancja prawdopodobieństw sukcesu (nierówność szans) mogła być większa lub mniejsza, np. (0,2;0,2; 0,8; 0,8) vs (0,2;0,4; 0,6; 0,8)
- Design pozwala odróżnić wpływ równości szans od altruizmu i stosunku do ryzyka

Powtórzenia

- Ilu wyborów mają dokonywać badani?
- Przykład: Dylemat Więźnia grany raz, vs grany kilkakrotnie, z różnymi partnerami
- Oczywiście, na ogół nie ma sensu powtarzać *tego samego pytania*
- Powtórzenia mogą mieć sens jeśli
 - Decyzje się różnią, np. wybór ceny zależnie od popytu
 - Pojawiają się nowe informacje, np. na temat wyników poprzednich decyzji, wyborów innych osób w analogicznej sytuacji (badania nad zachowaniami stadnymi itp.), wyborów partnerów interakcji strategicznej itd. itp.

One-shot vs powtórzenia

- **Pojedyncza decyzja:**
 - + Wyższa stawka w ramach danego budżetu
 - + Krótsze sesje
- **Powtórzenia**
 - + Więcej obserwacji
 - + Efekty uczenia się (ważne zwłaszcza w skomplikowanych designach)
 - + Reakcje na zmiany, np. w zachowaniu innego gracza
 - + Dynamika, np. konwergencja do równowagi

Schemat wypłaty za losowo wybraną rundę

- Na koniec eksper. losujemy jedną z rund
- Jedynie wynik w niej osiągnięty się liczy
- Badani mają zachętę by zawsze się starać
- Rzecz jasna *wartość oczekiwana* wypłaty mniejsza
- (np. przy 20 rundach mamy jedynie 5% szans, że dana runda będzie grała rolę, więc jeśli nominalna stawka decyzyjna wynosi 100 zł., to efektywna tylko 5 zł.)

Metoda strategii

- Sposób uzyskania większej liczby obserwacji od każdego uczestnika
- Zaproponowany przez Reinharda Seltena
- Badany podejmuje kilka lub kilkanaście decyzji wariantowych, z których tylko jedna (ale nie wiadomo z góry która) będzie miała wpływ na wyniki

Przykład zastosowania metody strategii

- W grze przetargu ultymatywnego Osoba 1 ma 10 zł. do podziału i wybiera ofertę dla Osoby 2, $x \in (0, 1, 2 \dots 10)$ zł. Druga osoba, zanim pozna x , wybiera opcję Przyjąć lub Odrzucić dla każdej z możliwych ofert.

Zalety metody strategii

- + Pozwala uzyskać więcej informacji naraz
- + Pozwala uzyskać reakcję nawet na bardzo rzadko wybieraną akcję pierwszego gracza

Wady metody strategii

- **Komplikuje zadanie gracza**
- **Zmniejsza stawki dla każdej decyzji, w porównaniu z jednokrotnym rozegranie danej gry**
- **Nie ma pewności, czy przynosi takie same wyniki jak tradycyjny sposób rozgrywki**
- **Np. reakcja może być „chłodniejsza” (*hot-cold empathy gap*)**
- **Badania (np. Brandts, Charness, 2000) na ogół nie potwierdzają jednak systematycznej różnicy w zachowaniu**

Matching

- **Jak tworzyć grupy w eksperymencie z powtarzanymi interakcjami?**
- **(Często ma to istotny wpływ na zachowanie. Inaczej będziemy postępować gdy spodziewamy się kolejnych interakcji z daną osobą.)**

Sposoby łączenia w grupy

- *Partner* (PM) – w kolejnych rundach badani pozostają w tych samych grupach
- *Stranger* (SM) – w każdej rundzie nowe losowanie, badani nie wiedzą czy grają ponownie z tą samą osobą
- *Perfect-stranger* (PSM) – w każdej rundzie nowe losowanie, dwóch badanych nie więcej niż raz w tej samej grupie
- *No-contagion* (NCM) – jak wyżej, przy czym dodatkowo dla każdego dwóch badanych A i B, którzy w którejkolwiek rundzie są w tej samej grupie, nie istnieje badany C, który byłby w którejś z rund w grupie z A a w innej rundzie w grupie z B.

Identyfikowanie wpływu zmiennych wyjaśniających

- Najczęściej chcemy zbadać jak zachowania zależnie od zmiennej albo zmiennych, które aktywnie manipulujemy (zmienne eksperymentalne)
- Zmienne te mają stanowić operacjonalizację pojęć teoretycznych lub nawiązywać do interesujących nas aspektów danej sytuacji społecznej
- Zmienna eksperymentalna przyjmie jeden z kilku „poziomów”, np. jeden w niektórych sesjach a drugi w innych
- Porównanie tendencji centralnej zmiennej zależnej pozwoli zidentyfikować efekt eksperymentalny
- Czasem naturalne jest traktować jeden z poziomów jako warunek kontrolny (*control treatment*, WK) a drugi „eksperymentalny” (WE)
 - (np. pewne oceny jako WK a te same oceny z systemem zachęt do trafności jako WE)

Wybór warunku kontrolnego

- Winien różnić się od WE w jeden sposób, wskazany przez teorię
- Przykład: *overconfidence*
- WE: uczestnicy decydują wziąć udział w teście kompetencji, w którym najlepsi otrzymają ustalone wynagrodzenie, np. 20 zł. a pozostali nic, czy też poprzestać na mniejszym lecz pewnym wynagrodzeniu, np. 10 zł.
- Możliwy WK: wybór pomiędzy pewną wypłatą (10 zł.) a loterią, zależną wyłącznie od szczęścia.
- Test kompetencji powinien, dla typowego uczestnika, dawać szansę sukcesu taką samą, co loteria z warunku kontrolnego
- Problem zmiennych zaburzających (*confounds*). Np. badani mogą chętnie wybrać test, sądząc, że będzie ciekawy albo chcąc się sprawdzić

Nie-samolubne preferencje jako zmienna zaburzająca

- **Przykład: współpraca w Dylemacie Więźnia powtarzanym skończoną liczbę razy**
- **Mogłaby posłużyć do falsyfikacji przewidywań teoriogrowych (opartych na tzw. indukcji wstecznej)**
- **Ale może być po prostu wyrazem altruizmu albo gotowości do warunkowej (odwzajemnionej) współpracy**

Wybór poziomów zmiennej eksperymentalnej

- Ile i jakich poziomów?
- *Default*: 2 poziomy, na tyle odległe, by wierzyć, że zrobi to różnicę
- Wybór poziomów winien być motywowany teorią (W szczególności, gdy mamy dwie alternatywne teorie, ustalmy poziomy zmiennej eksperymentalnej tak, by jedna przewidywała różnicę a druga jej brak)
- Dobrze by poziomy były „naturalne” lub skalibrowane z obserwacjami
- Należy unikać efektów podłogi i sufitu (bazowy poziom zmiennej zależnej na tyle „skrajny”, że nie ma już miejsca na efekt manipulacji eksperymentalnej)

Przykład: wybór poziomów zmiennej w Krawczyk (2010)

- Dwa poziomy koszty redystrybucji: 0 i 0,3
- Pierwszy jest naturalnym *benchmarkiem*
- Drugi nawiązuje do wielkości estymowanych w badaniach empirycznych.
- Oba na tyle niskie aby spodziewać się pewnego poziomu redystrybucji (uniknięcie efektu podłogi)
- Są też na tyle wzajemnie odległe aby przewidywania teoretyczne były zróżnicowane

Więcej niż dwa poziomy

- **Główny powód: oczekiwanie nieliniowości efektu.**
- **Np. model preferencji społecznych Fehra-Schmidta przewiduje, że wzrost relatywnego bogactwa da efekt nieróżniczkowalny w punkcie pełnej równości względem naszej grupy odniesienia.**

Randomizacja w obrębie sesji

- **Różni badani w tej samej sesji napotykają różne warunki**
 - **Bardziej kłopotliwe przygotowanie i prowadzenie sesji**
 - + **Już po pierwszej sesji mamy wstępne porównanie**
 - + **Kontrolujemy efekty czasowe (sesje poranne vs. wieczorne)**

Within-subject design

- (Ci sami badani najpierw podejmują decyzje w jednych warunkach, potem drugich)
 - + Te same zalety co na poprzednim slajdzie
 - + Większa moc testów (albo ta sama przy mniejszej liczbie uczestników)
 - + Możliwość identyfikacji konkretnych typów
 - Dłuższe sesje
 - Efekt dobrego badanego

Efekty czasowe

- **Efekt znudzenia**
- **Efekt uczenia się**
- *Spillover effect*

Technikalia

- **Laboratorium vs. Internet vs. kartka**
- **Komputer**
 - Ułatwia interakcje i ustrukturyzowaną komunikację
 - Umożliwia śledzenie procesu podejmowania decyzji (np. Mouselab)
 - Obniża koszty przetwarzania danych
- **Kartka+długopis**
 - Czasem łatwiejsze dla dużych grup
 - Być może większa wiarygodność

Próba

- **Zalety studentów**
 - **Niski koszt**
 - **Dostępność**
 - **Rozumienie instrukcji**
- **Zalecana ostrożność w przenoszeniu wyników ilościowych na całą populację**

Liczebność próby

- **Próba winna być tym liczniejsza, im**
 - **więcej jest warunków (grup) eksperymentalnych:**
 - **bardziej złożone są hipotezy (np. dotyczące efektu interakcji między zmiennymi)**
 - **słabszego efektu eksperymentalnego się spodziewamy**
 - **większy jest spodziewany szum w danych**
 - **mniej innowacyjny jest plan badawczy**
 - **bardziej ambitne są nasze zamierzenia publikacyjne**

Eksperymenty terenowe

- **Naturalność środowiska i decyzji**
- **Charakter dóbr**
- **Charakter informacji dostępnych badanym**
- **Wielkość stawek**
- **Populacja badanych**