




WOLAŃSKI



EWALUACJA EFEKTÓW WSPARCIA W RAMACH III OSI PRIORYTETOWEJ PONADREGIONALNA INFRASTRUKTURA KOLEJOWA PO PW 2014-2020, W TYM WPŁYWU NA REDUKCJĘ BARIER ROZWOJU MAKROREGIONU POLSKI WSCHODNIEJ

ZAŁĄCZNIK 2. METODYKA ANALIZ
KONTRFAKTYCZNYCH

SIERPIEŃ 2019



dr Michał Wolański

Paulina Kozłowska

Wiktor Mrozowski

Maciej Pańczak

Mateusz Pieróg

Łukasz Widła-Domaradzki

WSPÓLPRACA:

Jakub Kaczorowski

Badanie ewaluacyjne współfinansowane przez Unię Europejską w ramach Funduszu Spójności.

SPIS TREŚCI

1.	Analiza metodą PSM	2
1.1.	Podstawa teoretyczna.....	2
1.2.	Zastosowanie	3
2.	Analiza metodą SPSM	5
2.1.	Podstawy teoretyczne.....	5
2.2.	Zastosowanie	6

1. ANALIZA METODĄ PSM

1.1. PODSTAWA TEORETYCZNA

W ramach ewaluacji wpływu (ang. *impact evaluation*), która co do zasady ma na celu określenie zakresu oddziaływania interwencji publicznych, dostarczenie skwantyfikowanych dowodów możliwe jest jedynie przy zastosowaniu dwóch planów badawczych (ang. *research designs*): studium statystycznego (inaczej: szacowania strukturalnego) oraz schematu eksperymentalnego i quasi-eksperymentalnego. Potwierdzone zostało to w zaleceniach dotyczących metod badania wpływu pomocy publicznej, w których Komisja Europejska plany te określa jako dające „możliwość określenia przyczynowego wpływu samego programu, niezakłócanego przez inne zmienne, które mogły mieć wpływ na zaobserwowane wyniki, np. ogólne warunki makroekonomiczne (...)”¹.

Często stosowanym w ewaluacjach wpływu podejściem badawczym są eksperymenty i quasi-eksperymenty, które pozwalają na ustalenie efektów konkretnej interwencji poprzez wyznaczenie, a następnie porównanie (w ramach tej samej populacji) jednostek, które różnią się tylko jedną rzeczą - tym, że zostały lub nie zostały poddane interwencji. Co do zasady, podejście to cechuje się silną trafnością wewnętrzną (ang. *internal validity*), co oznacza, że zaobserwowane różnice w grupach eksperymentalnej i kontrolnej z dużym prawdopodobieństwem są wynikiem wdrożonej interwencji. W praktyce podejście czysto eksperymentalne jest niezwykle trudne do zastosowania, zwłaszcza na polu badania interwencji publicznych, dlatego też często stosuje się podejście quasi-eksperymentalne.

W literaturze przedmiotu zwykło się utożsamiać badanie efektu netto w ewaluacjach z użyciem metody *Propensity Score Matching*. Jest to metoda, którą w latach 80-tych XX wieku opracowano w celu wsparcia wyników badań eksperymentalnych. Chodziło m.in. o to, aby zniwelować efekt wpływu nielosowego doboru do grup eksperymentalnej i kontrolnej. Z punktu widzenia statystyka, metoda PSM ma swoje ograniczenia. Po pierwsze wymaga dobrze określonego i - najlepiej - dychotomicznie określonego efektu netto. Po drugie metoda dobierania „bliźniaka” lub „bliźniaków” do ewaluowanego obiektu wymaga zazwyczaj dużej bazy wyjściowej, z której dobierane są obiekty do parowania².

Pierwszym krokiem implementacji techniki PSM jest oszacowanie wartości *propensity score*. Na tym etapie należy podjąć decyzję, jaki wykorzystać model estymacji. Istnieją różnorodne metody szacowania $P(X_i)$ jednak najczęściej w literaturze wskazuje się na model logitowy lub probitowy - z preferencją dla tego pierwszego. Regresja logistyczna jest matematycznym modelem, który można wykorzystać do opisanie wpływu kilku zmiennych niezależnych x_1, x_2, \dots, x_k na dychotomiczną (przyjmującą wartość 0 lub 1) zmienną Y .

Model regresji logistycznej opiera się na funkcji logistycznej, która jest następującej postaci:

$$f(z) = \frac{e^z}{1+e^z}$$

Funkcja logistyczna przyjmuje wartości z przedziału (0; 1), dlatego może zostać wykorzystana do opisywania wartości prawdopodobieństwa. W szczególności może być to prawdopodobieństwo

¹ Komisja Europejska, Wspólne metody oceny pomocy państwa, Dokument Roboczy Służb Komisji, s. 7.

² Ł. Widła, D. Wojtowicz, M. Wolański i inni, Adaptacja metodologii pomiaru efektu netto interwencji publicznych do potrzeb sektora infrastruktury transportowej, Akademia Leona Koźmińskiego, Warszawa 2010

wzięcia udziału w pewnym zdarzeniu, np. programie infrastrukturalnym. Model regresji logistycznej wyraża się następującym równaniem:

$$P(D=1 | x_1, x_2, \dots, x_k) = \frac{e^{\beta_0 + \sum_{i=1}^k \beta_i x_i}}{1 + e^{\beta_0 + \sum_{i=1}^k \beta_i x_i}}$$

gdzie β_0 to stała regresji, β_i , $i = 1, \dots, k$ to współczynniki regresji, zaś x_1, x_2, \dots, x_k to zmienne niezależne, które mogą być zarówno ilościowe, jak i jakościowe. Aby oszacować β , wykorzystywana jest metoda największej wiarygodności (ang. *maximum likelihood*). Ogólna idea tej metody polega na szacowaniu wartości nieznanymi parametrów w taki sposób, aby te maksymalizowały prawdopodobieństwo uzyskania zaobserwowanych wartości zmiennej zależnej. Uzyskując wartości omawianego prawdopodobieństwa otrzymujemy również parametr określający podobieństwo między badanymi jednostkami (tu: odcinkami linii kolejowych).

W praktyce taka zbliżona wartość *propensity score* u dwóch różnych jednostek nie oznacza, że będą one miały takie same wartości na wszystkich zmiennych wykorzystanych w modelu. Możliwe jest, że będą się między sobą różnić wartościami wybranych zmiennych. Z różnych względów może się wydawać zasadne, aby w przypadku wybranych zmiennych w parach znajdowały się jednostki, które jedynie w zakresie pewnych charakterystyk są identyczne. Podejście takie znajduje szczególnie uzasadnienie w przypadku, gdy istnieją przypuszczenia co do tego, że oceniany efekt działania jest silnie zróżnicowany.

1.2. ZASTOSOWANIE

Metoda PSM dzieli wybraną populację (w tym przypadku odcinki linii kolejowych) na grupę eksperymentalną i kontrolną, czyli jednostki poddane i niepoddane interwencji, charakteryzujące się podobnym poziomem wejściowych wskaźników. W klasycznym podejściu metoda PSM bazuje na modelu regresji logistycznej, co oznacza, że warunkiem *sine qua non* jej wykorzystania jest zidentyfikowanie binarnej zmiennej zależnej. W przypadku omawianego badania założenie to jest spełnione, gdyż będzie możliwość określenia, które jednostki zostały poddane interwencji, a które nie. Zwykle w przypadku dopasowania zakłada się, że jednej jednostce eksperymentalnej (poddanej interwencji) przypisuje się jedną jednostkę kontrolną (niepoddaną interwencji) tzw. metodą najbliższego sąsiada (ang. *nearest neighbour*). Następnym krokiem jest określenie wysokości efektów netto dla jednostek poddanych interwencji (grupy eksperymentalnej) i jednostek niepoddanych interwencji (grupy kontrolnej).

W kolejnych krokach wykonuje się następujące czynności (por. Rysunek 1):

- dokonano podziału zbioru badanych jednostek na te poddane interwencji i pozostałe, na podstawie czego tworzy się zmienną binarną, gdzie wartość 1 oznacza przeprowadzenie interwencji;
- obliczenie *Propensity Score* dla wszystkich jednostek przy uwzględnieniu danych zmiennych;
- uszeregowanie odcinki według wartości *Propensity Score*;
- przyporządkowanie odcinka lub odcinków kontrfaktycznych o zbliżonej wartości *Propensity Score*.
- poprzez porównanie jednostek w ramach sytuacji kontrfaktycznych, oszacowany zostaje efekt netto interwencji.



Rysunek 1. Schemat postępowania w analizie netto metodą PSM



Lista zmiennych sukcesu, dla których szacujemy efekt netto interwencji, wynika z odtworzenia teorii interwencji. Obliczenie osiągniętych wartości, zestawienie ich z wartościami zakładanymi oraz poszukiwanie przyczyn takiego stanu rzeczy stanowi kluczowy element koncepcji ewaluacji opartej na teorii, jakim jest niniejsze badanie.

2. ANALIZA METODĄ SPSM

2.1. PODSTAWY TEORETYCZNE

Zastosowanie quasi-eksperymentalnego planu badawczego wykorzystującego metodę PSM w przypadku badania wpływu interwencji transportowych na wskaźniki gospodarcze na poziomie lokalnym ma jednak poważne ograniczenia. Po pierwsze, w klasycznym podejściu metoda PSM bazuje na modelu regresji logistycznej, co oznacza, że warunkiem *sine qua non* jej wykorzystania jest zidentyfikowanie binarnej zmiennej zależnej. W praktyce PSM można wykorzystywać w badaniu, w którym mamy do czynienia ze ściśle określonymi grupami jednostek poddanych i niepoddanych interwencji (w tym przypadku: poddanych i niepoddanych interwencjom infrastrukturalnym). Inną jest natomiast sytuacja, gdy wszystkie interesujące nas jednostki zostały poddane interwencji, tylko w różnym stopniu. Wykorzystanie regresji logistycznej nie jest więc możliwe w badaniu wpływu w sytuacji, w której każda jednostka terytorialna została poddana interwencji.

Analiza kontrfaktyczna wykorzystana w tym badaniu to metoda SPSM (*Stratified PSM*), pozwalająca na policzenie miar bliskości (*propensity scores*) dla powiatów.

Zaproponowana metoda opiera się na hierarchicznym algorytmie segmentowania Warda³, który jest metodą aglomeracyjną. Oznacza to, że każda obserwacja wchodzi do analizy jako pojedynczy segment, a następnie w kolejnych iteracjach najbliższe segmenty łączone są w jeden. Ostatecznie algorytm kończy pracę po uzyskaniu założonej uprzednio przez badacza liczby segmentów. Określanie optymalnej liczby segmentów w przypadku metod hierarchicznych wykracza poza zakres tego artykułu, istnieje jednak na ten temat bogata literatura⁴.

Algorytm Warda minimalizuje w każdym kroku błędy sumy kwadratów⁵ (ang. *error sum of squares* – *ESS*), w następujący sposób:

$$ESS = \sum_{i=1}^n x_i^2 - \frac{1}{n} (\sum_{i=1}^n x_i)^2 ,$$

gdzie x_i jest wartością dla i -tej jednostki, zaś n – wielkością próby. Algorytm powtarzany jest do momentu uzyskania ostatecznej liczby segmentów (lub „zbiorów”, wedle nomenklatury samego Warda).

Działanie każdej iteracji definiowane jest jako:

$$S(p_{n-1}, n - 1) = [S(p_{n-1}, n)] \cup [S(q_{n-1}, n)] ,$$

gdzie p_{n-1} jest mniejszym z dwóch segmentów (podzbiorów), zaś q_{n-1} jest większym z dwóch segmentów (zbiorów).

³ J. H. Ward, Hierarchical Grouping to Optimize an Objective Function, “American Statistical Association Journal” 1963.

⁴ D.G. Garson, Cluster Analysis, Statistical Associates Publishers 2014; C.C. Aggarwal, C.K. Reddy., Data Clustering: Algorithms and Applications, Chapman and Hall/CRC 2013; B.S. Everitt, S. Landau, M. Leese, D. Stahl, Cluster Analysis, Wiley 2011

⁵ Każdy krok, innymi słowy, ma za zadanie zmniejszać tzw. funkcję błędów i polepszać dopasowanie modelu do danych.

Procedura segmentowania hierarchicznego (podobnie jak inne procedury segmentowania, takie jak k-średnich czy k-median) oblicza domyślnie tylko przynależność do segmentu. W trakcie obliczania kolejnych kroków powstaje jednak również tablica dystansów - najczęściej w formie tablicy niepodobieństwa (ang. *dissimilarity matrix*) lub podobieństwa (ang. *similarity matrix*) - która może zostać użyta do obliczenia prawdopodobieństw przynależności do segmentów. Dzięki przekształceniom tablicy dystansów uzyskujemy wartości segmentowania rozmytego (ang. *fuzzy clustering*). Mając informację o przynależności do segmentu oraz dysponując odległościami z tablicy dystansów, można obliczyć rozmyte miary segmentowania hierarchicznego (ang. *fuzzy hierarchical clustering scores* - FHCS), które w następnym kroku posłużą jako miary PSM. FHCS definiowane jest jako:

$$FHCS_i = |DS_i - \bar{x}(DS)|,$$

gdzie DS to miara odległości (ang. *dissimilarity score*).

Procedura wyliczania miar segmentowania rozmytego została w niniejszym badaniu przeprowadzona dla każdego segmentu z osobna. Jak powiedziano powyżej, FHCS w dalszej procedurze traktowane były jako miary PSM, tak więc dalsza procedura uwzględniała różnice między wartościami każdego dwóch jednostek najbliższych leżących pod względem FHCS. Dla każdego segmentu można więc maksymalnie ustalić n-1 par, zakładając, że liczba jednostek w segmencie = n.

2.2. ZASTOSOWANIE

Jest to metoda analizy kontrfaktycznej umożliwiająca ustalenie efektów konkretnej interwencji poprzez identyfikację, a następnie porównanie jednostek, które różnią się nie faktem, lecz stopniem poddania interwencji. Podstawowym aspektem odróżniającym tę metodę od standardowej metody PSM (*Propensity Score Matching*) jest **podział jednostek na segmenty o podobnej charakterystyce przed rozpoczęciem interwencji oraz poszukiwanie podobnych do siebie jednostek w ramach tych segmentów**.

Metoda SPSM dzieli wybraną populację (w tym przypadku powiaty) na segmenty (klastry), charakteryzujące się podobnym poziomem wejściowych wskaźników. Innymi słowy, celem analizy jest podzielenie populacji na grupy maksymalnie różniące się między sobą oraz minimalnie różniące się wewnątrz klastra. Następny etap analizy polega na wyłonieniu wewnątrz klastra powiatów maksymalnie do siebie podobnych (oceniane jest to na podstawie miar PSM), zaś sama sytuacja kontrfaktyczna polega na porównaniu sytuacji par najbardziej podobnych do siebie powiatów. Porównując je ze sobą można uzyskać odpowiedź na pytanie o wpływ wskaźników transportowych na gospodarcze wskaźniki sukcesu.

W kolejnych krokach przeprowadzone zostaną następujące czynności (por. Rysunek 2):

- segmentacja jednostek na podstawie wartości wybranych zmiennych przed rozpoczęciem interwencji;
- za pomocą analizy dyskryminacyjnej⁶ w ramach segmentów zidentyfikowane zostaną jednostki najbardziej podobne do siebie, które w wyniku połączenia w pary, stworzą sytuacje kontrfaktyczne;

⁶ Metoda wielowymiarowej analizy danych polegająca m.in. na klasyfikacji badanych obiektów na podstawie określonych charakterystyk klas. W badaniu analiza dyskryminacyjna posłużyła do określenia prawdopodobieństwa przynależności miast do stworzonych segmentów w oparciu o wartości zmiennych

- porównując jednostki w ramach sytuacji kontrfaktycznych, oszacowany zostanie efekt netto interwencji;
- podjęta zostanie próba redukcji liczby zmiennych objaśniających za pomocą metod analizy czynnikowej⁷;
- przeprowadzona zostanie analiza regresji⁸ w celu identyfikacji oddziaływania transportowego, środowiskowego, społeczno-ekonomicznego oraz w obszarze bezpieczeństwa.

Rysunek 2. Schemat postępowania w analizie netto metodą SPSM



Pełna lista zmiennych, stosowanych w analizie tj. zmiennych objaśniających, objaśnianych i kontekstowych (służących do parowania – por. Rysunek 3) wynika z odtworzenia teorii interwencji, zaś ich późniejsza weryfikacja sprawdza, w jakim stopniu każda z nich oddziałuje na efekty, co jest realizacją postulatów TDE, czyli ewaluacji wspartej teorią (por. Rysunek 4).

segmentacyjnych. Rosnące uszeregowanie miast według tego prawdopodobieństwa pozwoliło na znalezienie jednostek najbardziej podobnych do siebie (sytuacji kontrfaktycznych).

⁷ Metoda wielowymiarowej analizy danych pozwalająca na zredukowanie liczby zmiennych do mniejszego zbioru w przypadku, gdy zmienne przenoszą podobny zakres informacji (np. gdyby analizowano jednocześnie liczbę zmodernizowanych jednostek taboru komunikacji miejskiej i liczbę zainstalowanych w nich wyświetlaczy).

⁸ Badanie związku między wielkościami danych, tj. między zmiennymi objaśnianymi a objaśniającymi.



Rysunek 3. Schemat wpływu na transportowe efekty interwencji



Rysunek 4. Teoria zmiany i jej weryfikacja

