

## STUDIA METODOLOGICZNE

### Od Redakcji

Zgodnie z zapowiedzią z poprzedniego numeru „Wiadomości Statystycznych”, tym razem zamieszczamy kolejne opracowania referatów z konferencji naukowej zorganizowanej przez Urząd Statystyczny w Szczecinie 23 i 24 marca br., pt. *Rola środowisk naukowych, samorządowych i służb statystyki publicznej we wzmacnianiu pozytywnego wizerunku statystyki*. Ich autorami są: Joanna Landmesser, Janusz Dygaszewicz i Agnieszka Nowakowska, Jan Purczyński i Rafał Klóska, Dorota Wyszowska oraz Hanna Dudek.

**Joanna LANDMESSER**

### Pomiar efektów oddziaływania różnych czynników na zmienne społeczno-ekonomiczne

---

Poprzez efekt oddziaływania danego czynnika na określone zjawisko rozumie się uśredniony wynik wpływu pewnej zmiennej losowej opisującej dwa potencjalnie stany — oddziaływania (*treatment*) i braku oddziaływania (*non-treatment*) owego czynnika na wyróżnioną zmienną endogeniczną. Termin „efekt oddziaływania” ma swoje źródło w literaturze medycznej dotyczącej badania skutków czynników, takich jak np. przeprowadzony w sposób nowatorski zabieg chirurgiczny. Obecnie pojęcie to znajduje zastosowanie w rozległych badaniach. Nowoczesne podejście do szacowania efektów oddziaływania pozwala odpo-

wiedzieć na różne pytania, np.: jakie są efekty zastosowania określonego nawozu w eksperymentach rolniczych? Jakie są skutki uczestnictwa w programie szkoleniowym dla szansy na zatrudnienie? Jakie są efekty netto różnego rodzaju świadczeń czy, ogólniej, polityki społecznej lub ekonomicznej?

W sferze badań ekonomicznych analizowane oddziaływanie może dotyczyć: programów pomocy dla bezrobotnych, programów edukacyjnych, różnego rodzaju regulacji prawnych, podjętej liberalizacji handlu międzynarodowego, przeprowadzonych interwencji czy realizowanej polityki wpływającej na otoczenie. Obiektami badania mogą być wtedy osoby fizyczne, gospodarstwa domowe, przedsiębiorstwa, a nawet kraje.

Opis zależności przyczynowo-skutkowych pomiędzy zachodzącymi zjawiskami może być przeprowadzony za pomocą różnych metod statystycznych. W naukach przyrodniczych często wykorzystywanym narzędziem są eksperymenty kontrolowane, pozwalające na analizę możliwych scenariuszy przebiegu zjawiska. W empirycznych badaniach ekonomicznych metody eksperymentalne stosuje się raczej rzadko ze względu na zachodzące ograniczenia techniczne, społeczne i etyczne (Strawiński, 2007). Wyizolowanie fragmentu gospodarki, tak aby spełnione były wymogi eksperymentu, jest trudne lub wręcz niemożliwe<sup>1</sup>. Ponadto rodzi się wątpliwość natury etycznej, czy wolno dla celów nauki eksperymentować na ludziach? Dodatkowo zachowanie osób biorących udział w eksperymencie mogłoby odbiegać od tego w warunkach rzeczywistych. W związku z czym, aby oszacować efekt oddziaływania czynnika, w sytuacji gdy nie można zastosować kontrolowanego eksperymentu zachodzi potrzeba wykorzystania odpowiednich metod statystycznych.

Celem artykułu jest przegląd metod statystycznych, stosowanych przy szacowaniu efektu oddziaływania czynników, wskazanie zalet i wad owych metod oraz zasygnalizowanie kierunków rozwoju w tym zakresie. Z danymi nieeksperymentalnymi wiąże się problem niewłaściwego doboru próby. Rozważona zostanie selekcja próby na podstawie czynników obserwowalnych oraz nieobserwowalnych. W pierwszym przypadku praktyczne znaczenie ma estymacja efektów oddziaływania czynników przez dopasowanie, w szczególności tzw. dobieganie na podstawie prawdopodobieństwa uczestnictwa (przedstawiony w opracowaniu przykład dotyczy właśnie tej metody). W drugim wskazano propozycje rozwiązań dla sytuacji, gdy ma miejsce selekcja na podstawie czynników nieobserwowalnych, jak estymator „różnic-w-różnicach”, estymacja metodą zmiennych instrumentalnych czy parametr *LATE*.

---

<sup>1</sup> Eksperyment zaczął funkcjonować w ekonomii jako uznana metoda badawcza dopiero wtedy, gdy w roku 2002 nagrodę Nobla otrzymali Vernon Smith i Daniel Kahneman. Doprowadziło to do intensywnego rozwoju gałęzi ekonomii behawioralnej. Obecnie w ekonomii wykorzystuje się najczęściej dwa rodzaje eksperymentów: laboratoryjne (krytykowane z powodu zbytnej „sterylności”) i terenowe (krytykowane za niepowtarzalność i niemożność przenoszenia wniosków na inne środowiska), Poskrobko (2012); Krawczyk (2012).

## ZDEFINIOWANIE EFEKTU ODDZIAŁYWANIA CZYNNIKA

Niech  $Y$  będzie symbolem badanej zmiennej (np. dochód, szansa na podjęcie pracy, ciśnienie krwi) poddawanej oddziaływaniu danego czynnika oraz  $D$  symbolem zmiennej binarnej opisującej elementy próby ze względu na poddanie oddziaływaniu:  $D = 1$  oznacza, że jednostkę poddano oddziaływaniu czynnika, natomiast  $D = 0$  oznacza, że jednostka nie została poddana jego oddziaływaniu (grupa kontrolna). Jeśli przez  $Y^1$  oznaczmy wartość zmiennej wynikowej wśród jednostek poddanych oddziaływaniu, a przez  $Y^0$  — wartość zmiennej wynikowej wśród jednostek niepoddanych oddziaływaniu określonego czynnika, to poszukiwanym efektem oddziaływania jest  $Y^1 - Y^0$ . Dla  $i$ -tej jednostki efekt ten możemy zdefiniować następująco  $Y_i^1 - Y_i^0$ . Niestety obserwowalne jest w danej chwili albo  $Y_i^1$ , albo  $Y_i^0$ , ale nigdy obie wartości razem.

Jeśli narażeni i nienarażeni na oddziaływanie danego czynnika stanowiliby próby losowe z populacji (np. grupa osób uczestniczących oraz grupa nieuczestniczących w szkoleniu byłyby próbami losowymi z populacji bezrobotnych), to poszukiwany efekt można by zmierzyć jako uśredniony efekt w próbie. Wtedy średni efekt oddziaływania (*average treatment effect*) można by było zdefiniować jako:

$$ATE = E(Y^1 - Y^0) = E(Y^1) - E(Y^0) = E(Y^1|D=1) - E(Y^0|D=0)$$

$ATE$  jest średnią różnicą między zmienną  $Y$  w grupie poddanej oddziaływaniu i w grupie niepoddanej oddziaływaniu czynnika (różnicą odpowiednich wartości oczekiwanych dla wyniku w grupach). Przy nielosowym doborze do próby estymator ten jest jednak obciążony.

Jeśli dobór do próby jest nielosowy, to nieobciążonym estymatorem dla efektu jest średni efekt oddziaływania czynnika na poddanych jego oddziaływaniu (*average treatment effect on the treated*):

$$ATT = E(Y^1 - Y^0|D=1) = E(Y^1|D=1) - E(Y^0|D=1)$$

$ATT$  jest różnicą między wartościami zmiennej  $Y$  poddanej i niepoddanej oddziaływaniu danego czynnika przy założeniu, że obie grupy zostały mu poddane. Składnik  $E(Y^0|D=1)$  tego równania jest nieobserwowalny i ma wartość tylko hipotetyczną (kontrafaktyczną). Odpowiada na pytanie, jaką wartość oczekiwaną przyjąłaby zmienna  $Y$  wśród osób poddanych oddziaływaniu, gdyby temu oddziaływaniu nie zostali poddani?

## METODY STATYSTYCZNE STOSOWANE PRZY SZACOWANIU EFEKTU ODDZIAŁYWANIA CZYNNIKA

Zalecenia co do stosowania odpowiednich metod statystycznych dla szacowania efektu oddziaływania czynnika uzależnione są od typu danych, jakimi dysponuje badacz — eksperymentalnych lub nieeksperymentalnych.

Źródłem danych eksperymentalnych są np. eksperymenty zrandomizowane. Randomizacja może być rozumiana jako proces, w wyniku którego rzut monety określa, czy np. osoba będzie leczona czy nie, tzn. podział na grupy poddanych i niepoddanych oddziaływaniu czynnika jest czysto losowy. Randomizacja zapewnia, że potencjalne efekty  $Y^1$  oraz  $Y^0$  są statystycznie niezależne od statusu oddziaływania. Wartość oczekiwaną  $E(Y^i)$  oblicza się jako średnią zaobserwowanych wartości  $Y^i$ . W takiej sytuacji zachodzi:

$$E(Y^1|D=1) = E(Y^1|D=0) = E(Y^1)$$

co skutkuje równością  $ATE = ATT$ . Niestety, jak już wcześniej sygnalizowano, eksperymenty są rzadkie w ekonomii, często zbyt drogie do wdrożenia, nie można ich uogólniać, a z ich realizacją łączy się wiele praktycznych trudności.

Warto wspomnieć, że dla danych eksperymentalnych celem oszacowania efektu oddziaływania czynnika możliwe jest zastosowanie podejścia regresyjnego. Wtedy estymator dla  $ATE$  stanowi parametr  $\beta_1$  w regresji  $Y_i = \beta_0 + \beta_1 D_i + \varepsilon_i$ .

Dysponując precyzyjną wiedzą dotyczącą zasad rządzących oddziaływaniem można skorzystać z koncepcji nieciągłości w regresji (*regression discontinuity*). Identyfikacja nieciągłości opiera się na obserwacji, że w świecie rządzonej przez zasady niektóre z praw są ustalane arbitralnie. Stanowią wtedy niejako przykłady dobrych eksperymentów. Z sytuacją taką można się spotkać, gdy istnieje próg dochodu ( $x_0$ ), przy którym ma miejsce kwalifikacja do pomocy finansowej lub próg wieku, w którym ktoś kwalifikuje się do planu emerytalnego.

Wtedy zmienna  $D_i$  jest zdefiniowana następująco  $D_i = \begin{cases} 1 & \text{gdy } x_i \geq x_0 \\ 0 & \text{gdy } x_i < x_0 \end{cases}$ .

Jak już zauważono, w wypadku danych nieeksperymentalnych pojawia się problem niewłaściwego doboru próby. Mówi się, że jednostki poddane oddziaływaniu określonego czynnika podlegają tzw. samoselekcji (LaLonde, 1986; Dehejia, Wahba, 1999; Morgan, Winship, 2007). Występuje ona, gdy badane jednostki same decydują czy chcą znaleźć się w grupie narażonej na oddziaływanie, czy też nie, np. dokonując wyboru pracując — nie pracując. Praca z danymi nieeksperymentalnymi oznacza brak spełnienia założenia niezależności zmiennej wynikowej od statusu oddziaływania:

$$\sim (Y^0, Y^1 \perp D)$$

Nawiązując do przykładu danych nieeksperymentalnych, dotyczących doboru osób na szkolenia dla bezrobotnych, zazwyczaj uczestniczący i nieuczestniczący w szkoleniu znacznie się różnią cechami od siebie. Gdyby osoby z wyższym wykształceniem częściej wybierały szkolenia dla bezrobotnych niż te z niższym, to można by przypuszczać, że efekty  $Y^1$  oraz  $Y^0$  będą skorelowane z wykształceniem. Zagadnienie to można interpretować jako problem pominiętych zmiennych.

*EFEKT ODDZIAŁYWANIA CZYNNIKA  
DLA SELEKCJI NA PODSTAWIE CZYNNIKÓW OBSERWOWALNYCH*

Chcąc poprawnie oszacować efekt oddziaływania czynnika należy przyjąć pewne dodatkowe założenie. Zakładamy, że dysponujemy wszystkimi zmiennymi  $X$  mającymi wpływ na badaną zmienną  $Y$  (co rozwiązuje problem pominiętych zmiennych)<sup>2</sup>. Formułujemy założenie warunkowej niezależności (*Conditional Independence Assumption* — CIA):

$$Y^0, Y^1 \perp D | X$$

Oznacza ono, że zmienna wynikowa  $Y$  jest warunkowo niezależna od doboru do grupy poddanych oddziaływaniu danego czynnika względem wektora zmiennych objaśniających  $X$ . Inaczej, przy ustalonym wektorze zmiennych  $X$  udział w tej grupie nie zależy od wartości zmiennej  $Y$  (Rubin, 1978; Cameron, Trivedi, 2005). Założenie CIA w zapisie dla wartości oczekiwanych przyjmuje postać  $E(Y^i | X, D) = E(Y^i | X)$ . Wtedy:

$$ATT(X) = E(Y^1 | X, D = 1) - E(Y^0 | X, D = 1) = E(Y^1 | X) - E(Y^0 | X) = ATE(X)$$

Niech  $E(Y^1 | X) = r_1(X)$  oraz  $E(Y^0 | X) = r_0(X)$ , gdzie  $r_1(\cdot)$  oraz  $r_0(\cdot)$  są funkcjami oddającymi zależność między wektorem  $X$  i oczekiwanymi wartościami odpowiednio dla  $Y^1$  oraz  $Y^0$ . Zgodny estymator dla średniego efektu oddziaływania  $ATE$  jest postaci:

$$\hat{ATE} = \frac{1}{N} \sum_{i=1}^N [\hat{r}_1(X_i) - \hat{r}_0(X_i)]$$

---

<sup>2</sup> Założenie to jest jednak mało realistyczne, ponieważ w praktyce prowadząc badanie nigdy do końca nie wiadomo, co wpływa na dane zjawisko oraz nie wszystkie, nawet znane, oddziałujące nań czynniki są mierzalne.

Jest to estymator bezwarunkowy, ponieważ ma tu miejsce uśrednianie po wszystkich jednostkach z różnymi wartościami wektora  $X$ . Z kolei estymator dla średniego efektu oddziaływania danego czynnika na poddanych jego oddziaływaniu ( $ATT$ ) przedstawia się jako:

$$\hat{ATT} = \frac{1}{N_T} \sum_{i=1}^{N_T} [\hat{r}_1(X_i) - \hat{r}_0(X_i)],$$

gdzie indeksy  $i = 1, \dots, N_T$  dotyczą tylko poddanych zabiegowi oddziaływania. Wobec powyższego  $\hat{ATE} \neq \hat{ATT}$ .

Efekty oddziaływania  $ATT$  mogą być oszacowane poprzez porównanie ze sobą osób z grupy poddanej oddziaływaniu oraz kontrolnej, dla których wartości  $X$  są identyczne. Postępowanie takie znane jest jako dopasowanie dokładne (*exact matching*). Brak w zbiorze danych obserwacji dla osób niepoddanych oddziaływaniu, które mają dokładnie te same wartości  $X$  jak osoba poddana, sprawia, że trudniej jest oszacować scenariusz kontrfaktyczny (np. gdy wektor  $X$  zawiera zmienne ciągłe). W praktyce stosowane jest wtedy dopasowanie niedokładne (dla wartości zbliżonych, ang. *inexact matching*). Dopasowanie może być szczególnie trudne do wykonania, gdy wymiar wektora  $X$  jest duży.

Rosenbaum i Rubin (1983) zaproponowali, aby w wypadku wielowymiarowego wektora  $X$  założenie warunkowej niezależności realizować na podstawie pewnego prawdopodobieństwa (metoda o nazwie ang. *propensity scores matching*, PSM). Miara *propensity score* oznacza prawdopodobieństwo uczestniczenia w grupie poddanych oddziaływaniu czynnika dla osoby opisanej przez wektor  $X$ , słowem  $\Pr(D=1|X) = p(X)$ . Założenie warunkowej niezależności rozszerza się w takim przypadku następująco:

$$Y^0, Y^1 \perp D | p(X)$$

wtedy  $ATE(p(X)) = E(Y^1|p(X)) - E(Y^0|p(X))$ .

Do ustalenia prawdopodobieństwa przynależności do grupy poddanych oddziaływaniu danego czynnika wykorzystuje się najczęściej model probitowy lub logitowy. Aby oszacować wartość kontrfaktyczną  $Y_i^0$  należy wykorzystać jedną lub kilka obserwacji z grupy kontrolnej, które są podobne do osoby  $i$ -tej w sensie *propensity score* (Heckman i in., 1997; Rosenbaum, Rubin, 1985). Dokonując estymacji przez dopasowanie wymaga się spełnienia założenia warunkowej

niezależności (CIA) oraz zachodzenia na siebie rozkładów dla  $X$  w grupie podanych oddziaływaniu czynnika i grupie kontrolnej (*common support*) —  $0 < \Pr(D = 1|X) < 1$ .

Estymator  $ATT$  w metodzie dopasowania przyjmuje postać:

$$ATT^M = \frac{1}{N_T} \sum_{i \in \{k: D_k=1\}} \left( Y_i^1 - \sum_{j \in \{k: D_k=0\}} \phi(i, j) Y_j^0 \right)$$

Symbol  $\phi(i, j)$  oznacza wagę przypisywaną dopasowanym obserwacjom. Wartość kontrfaktyczna jest wyznaczana jako średnia ważona wyników z grupy kontrolnej.

Popularną metodą łączenia obserwacji jest metoda najbliższego sąsiedztwa (*nearest neighbour matching*). Dla każdej szkolonej osoby określany jest zbiór  $A_i(X) = \{j | \min_j d(X_i, X_j)\}$ , gdzie  $d(X_i, X_j)$  oznacza odległość pomiędzy wektorami  $X_i$  oraz  $X_j$  (np. odległość Euklidesa lub Mahalanobisa). Jeśli dobieranie jest dokonywane za pomocą *propensity scores*, to zbiorem porównawczym jest  $A_i(p(X)) = \{p_j | \min_j d(p_i, p_j)\}$ . Możliwe jest dobieranie „jeden do jednego” (wtedy  $\phi(i, j) = 1$ ) lub „jeden do kilku”.

Inna metoda, suwmiarki (*caliper matching*), polega na nałożeniu na odległości  $d(p_i, p_j)$  górnego ograniczenia, utrzymując warunek określonej liczby połączeń. Z kolei metoda promienia (*radius matching*) dopuszcza wszystkie połączenia nieprzekraczające pewnej z góry ustalonej różnicy między prawdopodobieństwami. Istotą metody przedziałowej (*interval matching*) jest podział próby na przedziały, dla których  $\Pr(D = 1)$  jest w obu podgrupach zbliżone. Efekt oddziaływania czynnika oblicza się osobno dla każdej pary przedziałów, a efekt dla całej próby jest ważoną średnią oszacowań cząstkowych. W popularnej metodzie estymacji jądrowej (*kernel matching*) wagi  $\phi(i, j)$  są zdefiniowane z wykorzystaniem funkcji jądra  $K(\cdot)$ , czyli symetrycznej funkcji gęstości, osiągnącej maksimum dla  $p_i = p_j$  (wtedy argument  $K(\cdot)$  wynosi zero) i wartościach zmniejszających się w miarę wzrostu wartości bezwzględnej argumentu.

Zaletą metody dopasowania jest fakt, że przeprowadzona w ten sposób estymacja efektu oddziaływania czynnika, w przeciwieństwie do podejścia regresyjnego, nie zakłada zależności funkcyjnej między oddziaływaniem owego czynnika a uzyskanym wynikiem. Jest to więc metoda nieparametryczna. Dodatkowo, w wypadku estymacji przez dopasowanie nie przeprowadza się ekstrapolacji. Jednak brak możliwości oszacowania efektu oddziaływania czynnika poza zakresem *common support* stanowi wadę metody. Wariancję estymatora  $ATT$ , uzyskanego za pomocą prawdopodobieństwa PSM, można oszacować jedynie za pomocą metod przybliżonych, w szczególności metod bootstrapowych.

*EFEKT ODDZIAŁYWANIA CZYNNIKA  
DLA SELEKCJI NA PODSTAWIE CZYNNIKÓW NIEOBSERWOWALNYCH*

Może się zdarzyć, że selekcja do próby odbywa się na podstawie czynników nieobserwowalnych. W takim wypadku założenie warunkowej niezależności  $Y^0, Y^1 \perp D|X$  nie jest spełnione, ponieważ występuje nieobserwowalny czynnik wpływający zarówno na status oddziaływania czynnika  $D$ , jak i potencjalny wynik  $Y$ .

Gdy taka nieobserwowalna zmienna jest stała, można wykorzystać dane wzdłużne (panelowe) i usunąć ten czynnik przez różnicowanie. Propozycją dla estymatora  $ATT$  w takiej sytuacji jest estymator „różnic-w-różnicach” (*differences-in-differences estimator*) (Card, Krueger, 1994):

$$ATT = (\bar{Y}_a^T - \bar{Y}_b^T) - (\bar{Y}_a^{nT} - \bar{Y}_b^{nT})$$

Subskrypty  $a$  i  $b$  oznaczają odpowiednio „after” i „before”, czyli po i przed eksperymentem. Estymator ten jest zdefiniowany jako różnica w zmianach uśrednionych wartości wynikowych dla grupy poddanej oddziaływaniu czynnika ( $T$ ) oraz grupy kontrolnej ( $nT$ ).

Alternatywną metodą przy identyfikacji efektu oddziaływania czynnika, gdy ma miejsce selekcja na podstawie czynników nieobserwowalnych, jest metoda zmiennych instrumentalnych (IV estymacja) (Angrist i in., 1996).

Rozważmy model liniowy  $Y_i = \beta X + \alpha D + u$ . Składnik losowy  $u$  obejmuje pominięte obserwowalne i nieobserwowalne czynniki wpływające na zmienną  $Y$ . Jeśli ma miejsce korelacja pomiędzy zmienną  $D$  oraz składnikiem losowym  $u$ , to estymacja modelu KMNK (klasyczna metoda najmniejszych kwadratów) dostarczy obciążonych ocen parametrów. W metodzie IV zmienna  $D$  jest zastępowana egzogenicznym instrumentem  $Z$ , który jest skorelowany z  $D$ , ale nie jest skorelowany z  $u$ . Estymacja IV polega więc w pewnym sensie na symulowaniu randomizacji. Niewątpliwą wadą tej metody jest trudność w znalezieniu odpowiedniego instrumentu w ramach rozwiązywania problemów z zakresu nauk społecznych.

Innym miernikiem wykorzystywanym do oceny efektów oddziaływania czynnika, estymowanym za pomocą metody zmiennych instrumentalnych, jest *LATE* (*local average treatment effect*). Angrist (1990) rozważał wpływ statusu weterana (zmienna  $D$ ) na wynagrodzenia w Stanach Zjednoczonych (zmienna  $Y$ ). W 1970 r., aby uniknąć podejrzeń o manipulacje przy powołaniach do służby wojskowej, wprowadzono loterię, która wyłaniała pulę potencjalnych powołanych. Część wyłonionych była jednak niezdolna do służby wojskowej, podczas gdy inni niewyłonieni stali się ochotnikami.

Niech  $Z_i = 1$ , jeśli loteria kwalifikowała do służby, a  $Z_i = 0$ , jeśli loteria nie kwalifikowała. Dodatkowo, niech  $D_i = 1$ , jeśli osoba służyła w Wietnamie oraz

$D_i = 0$ , jeśli osoba nie służyła. Zmienna  $Z$  jest instrumentem skorelowanym ze statusem weterana. Zmienna  $Z$  wywiera wpływ na zmienną  $D$ , ta zaś oddziałuje na zmienną  $Y$  ( $Z_i \rightarrow D_i \rightarrow Y_i$ ). Estymator *LATE* wyraża średni efekt oddziaływania czynnika tylko w subpopulacji „pokornych”, czyli takich, którzy zachowali się zgodnie ze wskazaniem loterii (inne możliwe subpopulacje to: *never-takers* — niesłużący niezależnie od wyniku loterii; *always-takers* — służący niezależnie od wyniku loterii; *defiers* — zawsze się sprzeciwiający) (Imbens, Angrist, 1994):

$$LATE = E(Y_i^1 - Y_i^0 | D_i^1 - D_i^0 > 0)$$

We wzorze tym zmienna  $D_i^1 = 1$  wtedy, gdy daną osobę zakwalifikowano do służby w wyniku loterii na wyjazd do Wietnamu, natomiast  $D_i^1 = 0$ , gdy pomimo wskazania przez loterię do służby, osoba ta nie służyła w Wietnamie. Podobnie  $D_i^0 = 1$ , gdy dana osoba nie została zakwalifikowana w loterii, ale mimo to służyła w Wietnamie oraz  $D_i^0 = 0$ , przy braku wskazania w loterii do służby i braku wyjazdu.

Szczególne znaczenie *LATE* wynika stąd, że umożliwia on określenie rozmiaru efektu oddziaływania czynnika w badaniach łączących się z licznymi odmowami uczestnictwa. *LATE* pozwala na prawidłowe zaprojektowanie badań, w których wymuszenie uczestnictwa jest niemożliwe.

### EMPIRYCZNA ILUSTRACJA ZAGADNIENIA

Niech ilustracją właściwego pomiaru efektu oddziaływania określonego czynnika będzie ocena skuteczności szkoleń dla bezrobotnych, przeprowadzona na podstawie danych uzyskanych z Powiatowego Urzędu Pracy w Słupsku z lat 2000—2007. Dane dostępne w bazie PUP nie stanowią wyniku eksperymentu (wybór osób do udziału w szkoleniach nie odbywał się w sposób losowy), czyli są przykładem danych nieeksperymentalnych. Każda analiza efektu oddziaływania szkoleń narażona jest na popełnienie ewentualnych błędów we wnioskowaniu, wynikających z nielosowego doboru próby.

Skuteczność szkolenia można ocenić posługując się modelem hazardu Coxa postaci  $h(t; X_{it}, D_{it}) = h_0(t) \exp(\beta X_{it} + \alpha D_{it})$ , w którym  $h_0(t)$  to tzw. hazard bazowy, a składowa  $\exp(\beta X_{it} + \alpha D_{it})$  charakteryzuje wpływ cech indywidualnych osób  $X$  oraz udziału w szkoleniu  $D$  na szansę zmiany stanu (Cox, 1972). Model Coxa stosuje się, gdy badacz jest zainteresowany wpływem zmiennych na czas trwania w pewnym stanie, mniej zwracając uwagę na zależność od czasu, a także gdy dane empiryczne charakteryzują się cenzurowaniem. Teoretycznie interesuje nas wyznaczenie różnicy pomiędzy wartością zmiennej wyniko-

wej określającej szansę na podjęcie zatrudnienia przez osobę szkoloną a wartością tej zmiennej dla tej samej osoby, ale bez udziału w szkoleniu. Ta druga sytuacja jest hipotetyczna.

W pierwszym kroku oszacowano model Coxa dla połączonej próby epizodów bezrobocia osób szkolonych (625 epizodów) oraz epizodów osób nigdy nieszkolonych (5573 epizody). Następnie za pomocą techniki z zakresu metod doborania skonstruowano odpowiednie grupy kontrolne dla grupy osób zaangażowanych w szkolenia. Celem ustalenia prawdopodobieństwa udziału w szkoleniu oszacowano model probitowy z takimi zmiennymi objaśniającymi, jak: wiek osoby, jej wykształcenie, miejsce zamieszkania oraz ewentualna niepełnosprawność (na temat doboru zmiennych pisali Heckman i in., 1998 oraz Hujer i in., 1997). Następnie wykonano: dobieranie *propensity score* „1-to-1”, oparte na algorytmie najbliższego sąsiedztwa (NN) ze zwracaniem; dobieranie *propensity score* „1-to-4”, NN ze zwracaniem oraz dobieranie dokładne „1-to-1”, oparte na odległościach Mahalanobisa (Rubin, 1980). Do każdego uczestnika szkolenia dobrano jedną osobę (ewentualnie cztery osoby) niebędącą uczestnikiem szkolenia, która była jednak do niego podobna pod względem indywidualnych cech (Heckman i in., 1997; Cameron, Trivedi, 2005). Dla tak skonstruowanych prób oszacowano modele Coxa dla czasu trwania w stanie bezrobocia.

#### WYNIKI ESTYMACJI MODELI HAZARDU COXA DLA SZANSY OPUSZCZENIA STANU BEZROBOCIA

Zmienne	Model bez doborania	Model z doboraniem <i>ps</i> 1-1 NN ze zwracaniem	Model z doboraniem <i>ps</i> 1-4 NN ze zwracaniem	Model z doboraniem 1-1 Mahalanobisa
	exp( $\beta_k$ )			
<i>pleć</i> .....	1,505 <sup>a</sup>	1,682 <sup>a</sup>	1,707 <sup>a</sup>	1,667 <sup>a</sup>
<i>wiek</i> .....	0,994 <sup>a</sup>	0,933 <sup>b</sup>	0,994 <sup>b</sup>	0,994 <sup>b</sup>
<i>wyższe</i> .....	1,686 <sup>a</sup>	2,422 <sup>a</sup>	1,865 <sup>a</sup>	2,374 <sup>a</sup>
<i>średnie</i> .....	1,319 <sup>a</sup>	1,474 <sup>a</sup>	1,258 <sup>a</sup>	1,520 <sup>a</sup>
<i>zawodowe</i> .....	1,172 <sup>a</sup>	1,675 <sup>a</sup>	1,135	1,709 <sup>a</sup>
<i>niesprawność</i> .....	0,687 <sup>a</sup>	0,575 <sup>a</sup>	0,650 <sup>a</sup>	0,536 <sup>a</sup>
<i>szkol_akt</i> .....	0,639 <sup>a</sup>	0,665 <sup>a</sup>	0,660 <sup>a</sup>	0,671 <sup>a</sup>
<i>szkol_pas</i> .....	1,412 <sup>a</sup>	1,591 <sup>a</sup>	1,531 <sup>a</sup>	1,614 <sup>a</sup>
liczba obserwacji .....	6198	1250	1319	1250
<i>lnL</i> .....	-41713,922	-6299,842	-6818,242	-6292,608

*a, b* Istotność statystyczna odpowiednio: *a* — 1%, *b* — 5%.

Źródło: obliczenia własne z wykorzystaniem procedury *psmatch2* w programie Stata.

Modele te zawierały m.in. dwie zmienne dychotomiczne: *szkol\_akt* — zmienna oznaczająca uczestnictwo w szkoleniu w trakcie aktualnego epizodu bezrobocia oraz *szkol\_past* — zmienna oznaczająca uczestnictwo w szkoleniu w ostatnich 3 latach przed rejestracją osoby w UP jako bezrobotna.

Można zaobserwować, że w badanej próbie uczestnictwo w szkoleniu w ciągu ostatnich 3 lat przed nową rejestracją w urzędzie pracy znacznie zwiększało prawdopodobieństwo krótszego czasu trwania kolejnego epizodu bezrobocia (tablica). Odpowiedni parametr w modelu bez dobierania wskazuje, że szanse podjęcia pracy przez osobę po szkoleniu były wyższe o 41% niż u osoby bez kursu. Efekt ten okazał się jednak wyższy o ok. 20% dla modeli oszacowanych na podstawie specjalnie dobranych prób. Z kolei oszacowania współczynników hazardu, wiążące się ze zmienną *szkol\_akt*, świadczą o negatywnym wpływie aktualnego uczestnictwa w szkoleniu na szansę opuszczenia stanu bezrobocia (szanse te spadały o ok. 35%). Działo się tak zapewne dlatego, że zaabsorbowani kursem bezrobotni nie prowadzili w tym samym czasie intensywnych poszukiwań pracy. W modelu opartym na niedobrej próbie efekt ten był przeszacowany (Landmesser, 2011a, 2011b).

Przedstawiony przykład analizy oddziaływania czynnika zakładał selekcję na podstawie czynników obserwowalnych. Przyjmując jednak, że selekcja do szkoleń odbywała się na podstawie czynników nieobserwowalnych, problem doboru próby można by rozwiązać instrumentalizując udział w szkoleniu. Zachodzi jednak niewątpliwa trudność w znalezieniu odpowiedniego instrumentu. Przykład podjętej próby w tym zakresie stanowi praca (Landmesser, 2010).

## Podsumowanie

Artykuł poświęcono przeglądowi metod statystycznych stosowanych przy szacowaniu efektu oddziaływania określonego czynnika. Zaprezentowano technikę możliwą do wykorzystania w badaniach ewaluacyjnych. Można spodziewać się, że ich popularność — szczególnie w badaniach ekonomicznych — będzie rosła. Zawsze jednak możliwość zastosowania odpowiedniej techniki będzie zależała od tego, jakimi danymi dysponuje badacz.

Kierunki rozwoju metod oceny efektów oddziaływania czynnika sprowadzają się do podejmowania prób uogólniania znanych metod i osłabiania założeń proponowanych modeli. Gruszczynski (2012) zwraca uwagę na odejście w konstruowanych współcześnie modelach od binarnego charakteru oddziaływania danego czynnika i dopuszczenie zróżnicowania intensywności jego oddziaływania. Podejście takie staje się atrakcyjne w wypadku chęci oceny udziału jednostek w więcej niż jednym programie. Istnieją również próby oszacowania rozkładu efektu oddziaływania czynnika w populacji (Aakvik i in., 2005).

Drugim nurtem badań nad estymatorami efektów oddziaływania czynnika są ich asymptotyczne własności. W przypadku niektórych metod problematyczne jest szacowanie wariancji tych estymatorów (np. *PSM*). Stosowane jednak w tym celu metody bootstrapowe wciąż spotykają się z krytyką (Abadie, Imbens, 2008).

## LITERATURA

- Aakvik A., Heckman J. J., Vytlacil E. (2005), *Estimating Treatment Effects for Discrete Outcomes When Responses to Treatment Vary: an Application to Norwegian Vocational Rehabilitation Programs*, „Journal of Econometrics”, No. 125
- Abadie A., Imbens G. W. (2008), *On the Failure of the Bootstrap for Matching Estimators*, „Econometrica”, No. 76
- Angrist J. D. (1990), *Lifetime Earnings and Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records*, „American Economic Review”, No. 80
- Angrist J. D., Imbens G. W., Rubin D. B. (1996), *Identification of Causal Effects Using Instrumental Variables*, „Journal of the American Statistical Association”, No. 91 (434)
- Cameron A. C., Trivedi P. K. (2005), *Microeconometrics: Methods and Applications*. Cambridge University Press, New York
- Card D., Krueger A. B. (1994), *Minimum Wages and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania*, „American Economic Review”, No. 84 (4)
- Cox D. R. (1972), *Regression Models and Life Tables (with Discussion)*, „Journal of the Royal Statistical Society”, Series B 34
- Dehejia R. H., Wahba S. (1999), *Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs*, „Journal of the American Statistical Association”, No. 94
- Gruszczyński M. (red.) (2012), *Mikroekonometria. Modele i metody analizy danych indywidualnych*, Wolters Kluwer Polska Sp. z o.o., Warszawa
- Heckman J. J., Ichimura H., Todd P. E. (1997), *Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme*, „Review of Economic Studies”, No. 64
- Heckman J. J., LaLonde R. J., Smith J. A. (1998), *The Economics and Econometrics of Active Labour Market Programmes*, [w:] Ashenfelter O., Card D. (wyd.), *The Handbook of Labour Economics*, Vol. 3A, Elsevier, New York
- Hujer R., Maurer K.-O., Wellner M. (1997), *The Impact of Training on Unemployment Duration in West Germany — Combining a Discrete Hazard Rate Model with Matching Techniques*, „Frankfurter Volkswirtschaftliche Diskussionsbeiträge”, No. 74, Johann Wolfgang Goethe-University, Frankfurt/Main
- Imbens G. W., Angrist J. (1994), *Identification and Estimation of Local Average Treatment Effect*, „Econometrica”, No. 62
- Krawczyk M. (red.) (2012), *Ekonomia eksperymentalna*, Wolters Kluwer Polska Sp. z o.o. Warszawa
- LaLonde R. J. (1986), *Evaluating the Econometric Evaluations of Training Programs*, „American Economic Review”, No. 76 (4)
- Landmesser J. (2010), *Assessing the Impact of Training on Unemployment Duration Using Hazard Models with Instrumental Variables*, „Quantitative Methods in Economics”, Vol. XI, No. 1
- Landmesser J. (2011a), *Mikroekonometryczne metody pomiaru skuteczności szkoleń dla bezrobotnych*, Taksonomia 18, Klasyfikacja i analiza danych — teoria i zastosowania (red. K. Jajuga, M. Walesiak), „Prace Naukowe UE we Wrocławiu”, nr 176
- Landmesser J. (2011b), *The Impact of Vocational Training on the Unemployment Duration*, „International Advances in Economic Research”, No. 17
- Morgan S., Winship C. (2007), *Counterfactuals and Causal Inference: Methods and Principles for Social Science*, Cambridge University Press, New York
- Poskrobko T. (2012), *Nowe koncepcje i teorie rozwoju w ekonomii*, [w:] Czajka S., Poskrobko T. i in., *Wyzwania współczesnej ekonomii*, Wydawnictwo Difin, Warszawa

- Rosenbaum P. R., Rubin D. B. (1983), *The Central Role of the Propensity Score in Observational Studies for Causal Effects*, „Biometrika”, nr 70 (1)
- Rosenbaum P. R., Rubin D. B. (1985), *Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score*, „The American Statistician”, No. 39 (1)
- Rubin D. B. (1978), *Bayesian Inference for Causal Effects*, „Annals of Statistics”, No. 6
- Rubin D. B. (1980), *Bias Reduction Using Mahalanobis-Metric Matching*, „Biometrics”, No. 36
- Strawiński P. (2007), *Przyczynowość, selekcja i endogeniczne oddziaływanie*, „Przegląd Statystyczny”, nr 4

## SUMMARY

*The aim of this article is an overview of modern statistical methods used in estimating the treatment effects, showing their advantages and disadvantages, and to indicate the directions of development. Due to the rarely performed in economic sciences randomized experiments, attention has been focused on the analysis of non-experimental data. The problem related to this type of data is inadequate sampling. We have considered the selection on observable and unobservable factors. The special importance of proper sample selection for the estimation of the treatment effects has been documented with an empirical example for evaluating the effectiveness of vocational training for the unemployed.*

*Special importance of proper matching attempts to estimate the effects of the impact of certain factors documented empirical example, for evaluating the effectiveness of the training for the unemployed.*

## РЕЗЮМЕ

*Целью статьи является анализ современных статистических методов используемых в оценке эффекта воздействия различных факторов на исследовательские переменные, представление их достоинств и недостатков, а также сигнализация направлений развития. Учитывая редко проводимые в экономических науках рандомизированные эксперименты, внимание было сосредоточено на анализе неэкспериментальных данных. К такого типа данным относится проблема неправильного подбора выборки. Рассматривался отбор выборки на основе наблюдаемых и ненаблюдаемых факторов. Особенное значение соответствующего подбора выборки для оценки эффектов воздействия некоторых факторов было подтверждено эмпирическим примером, касающимся оценки эффективности обучения безработных.*