

*Dominik Śliwicki**

ZASTOSOWANIE ESTYMATORÓW JĄDROWYCH DO SZACOWANIA EFEKTYWNOŚCI AKTYWNYCH PROGRAMÓW RYNKU PRACY

Z a r y s t r e ś c i. Każda interwencja prowadzona przez państwo na rynku pracy musi być po zakończeniu ewaluowana pod kątem efektywności. Celem artykułu jest ocena efektywności aktywnych programów rynku pracy. Jako narzędzie badawcze zostanie wykorzystana technika *Propensity Score Matching* oraz podejście oparte na estymatorach jądrowych. Rozważania teoretyczne zostaną poparte analizą empiryczną.

S ł o w a k l u c z o w e: efektywność netto programów rynku pracy, estymator jądrowy.

K l a s y f i k a c j a J E L: C13, C14, J68

WSTĘP

Współczesne wysoko rozwinięte państwa prowadzą w ramach polityki społeczno-gospodarczej aktywne działania na rynku pracy – zalicza się do nich wszystkie formy interwencji publicznej, których głównym celem jest zwiększenie szans na podjęcie i utrzymanie zatrudnienia przez osoby znajdujące się w trudnej sytuacji na rynku pracy oraz przeciwdziałanie ich dezaktywizacji, a także aktywizacja zawodowa osób biernych zawodowo. Każda podejmowana w ramach strategii na rynku pracy interwencja musi być po zakończeniu ewaluowana pod kątem skuteczności (Bieliński, Bober, Sarzańska, Zawistowski, 2008). Zasadniczym elementem ewaluacji *ex-post* jest oszacowanie efektu przyczynowego podjętego działania, tj. określenie, na ile

* Adres do korespondencji: Dominik Śliwicki, Urząd Statystyczny w Bydgoszczy, ul. Kornarskiego 1–3, 85-066 Bydgoszcz, e-mail: dominik_sliwicki@o2.pl.

osiągnięty rezultat jest efektem rzeczywistego wpływu zrealizowanej interwencji. Z metodologicznego punktu widzenia najodpowiedniejszą formą badania efektu przyczynowego byłby prawdziwy eksperyment. Jednakże jego wykorzystanie w naukach społecznych w znacznej większości przypadków nie jest możliwe z uwagi na to, że nie ma możliwości kontrolowania wszystkich czynników mogących wywierać wpływ na końcowy efekt (Konarski, Kotnarowski, 2007). W metodach eksperymentalnych z reguły występują dwie grupy osób: objętych oddziaływaniem programu oraz nieuczestniczących w programie. Alternatywę stanowią metody nieeksperymentalne (*quasi-eksperymentalne*), które koncentrują się na badaniu osób uczestniczących w aktywnych programach rynku pracy. W metodach tych nie występuje grupa odniesienia w postaci osób nie objętych działaniem interwencji (Maksim, 2008).

1. PODSTAWY EWALUACJI AKTYWNYCH PROGRAMÓW RYNKU PRACY

W praktyce gospodarczej najczęściej szacowane są efekty brutto aktywnych programów rynku pracy, oznaczające odsetek osób, które znalazły zatrudnienie po uczestnictwie w aktywnym programie rynku pracy. Oprócz samego efektu przyczynowego zastosowanej interwencji, efekt brutto zawiera również szereg efektów zewnętrznych, do których zaliczyć można m.in. efekt jałowego biegu oraz efekt spijania śmietanki. Efekt jałowego biegu występuje wtedy, gdy podejmowane działania nie przynoszą efektów zatrudnieniowych, np. bezrobotny, który uczestniczył w programie, mógłby podjąć zatrudnienie bez brania w nim udziału – w trakcie programu nie nabył nowych umiejętności. Efekt spijania śmietanki natomiast polega na tym, że do programu kierowane są osoby najlepiej rokujące na znalezienie pracy. Efekty zewnętrzne zamazują rzeczywisty wpływ przyczynowy aktywnego programu rynku pracy. W związku z tym powstała konieczność szacowania efektu netto programów rynku pracy.

Rozważania naukowe nad zagadnieniem przyczynowości w badaniach ewaluacyjnych doprowadziły do wypracowania przez statystyków i ekonometryków, wywodzącej się z logiki, koncepcji stanów kontrfaktycznych (*counterfactual framework*). Umożliwia ona operacyjne podejście do wnioskowania o wielkości efektu przyczynowego w sytuacjach nieeksperymentalnych. Podstawowe założenie koncepcji stanów kontrfaktycznych polega na tym, że osoby poddane działaniu określonego instrumentu aktywnej polityki rynku pracy oraz osoby nie poddane działaniu tego instrumentu posiadają potencjalne wyniki w obu tych stanach. Przy czym dana osoba może być

obserwowana tylko w jednej sytuacji, tzn. jako beneficjent formy interwencji (grupa badana) albo osoba nieuczestnicząca w formie interwencji (pula kontrolna). Rezultatem takiego podejścia jest to, że wyniki zaobserwowane dla osób badanych w jednym stanie (interwencji albo kontrolnym) mają potencjalne wyniki w drugim stanie (Konarski, Kotnarowski, 2008). Oznaczając przez Y_i^1 efekt dla osoby poddanej interwencji, a przez Y_i^0 potencjalny wynik w sytuacji braku uczestnictwa w interwencji, efekt przyczynowy interwencji (*treatment effect*) dla i -tej osoby można określić jako różnicę pomiędzy nimi, czyli (Caliendo, 2006):

$$\Delta_i = Y_i^1 - Y_i^0. \quad (1)$$

Wyznaczenie wprost indywidualnego efektu interwencji danego wzorem (1) jest niemożliwe z uwagi na to, że nie można dla tej samej osoby zaobserwować jednocześnie efektu jednej i drugiej sytuacji. Problem ten nosi nazwę fundamentalnego problemu wnioskowania przyczynowego (*fundamental problem of causal inference*). Koncentracja na pojedynczej osobie wymaga, aby efekt zastosowanej interwencji dla każdej osoby był niezależny od udziału w tej interwencji innych osób. Założenie to jest określane jako *stable unit treatment assumption* (SUTVA). Jeden ze składników formuły (1) jest nieobserwowalny i nazywa się go właśnie stanem kontrfaktycznym. W związku z tym obserwowany efekt dla i -tej osoby dany jest formułą (Caliendo, 2006)::

$$Y_i = D_i Y_i^1 - (1 - D_i) Y_i^0, \quad (2)$$

gdzie: D_i jest zmienną binarną przyjmującą wartość 1, gdy i -ta osoba należy do grupy badanej, oraz 0, gdy ta osoba należy do puli kontrolnej.

Jako rozwiązanie fundamentalnego problemu wnioskowania przyczynowego wskazuje się metody imputacji, tzn. pozyskania informacji o efekcie w stanie kontrfaktycznym od osób nieuczestniczących w formie interwencji – dla nich obserwowalny jest efekt braku interwencji. Ponieważ nigdy nie uda się wyznaczyć indywidualnego efektu dla konkretnej osoby, problem przenoszony jest na poziom populacji, z której ta osoba pochodzi. W tym kontekście rozpatruje się wartości średnie celów interwencji. Do najczęściej występujących w literaturze parametrów efektywności programów dla populacji należą:

- średni efekt interwencji (ATE – *average treatment effect*)

$$\Delta_{ATE} = E(\Delta) = E(Y^1) - E(Y^0), \quad (3)$$

- średni efekt interwencji dla osób poddanych interwencji (ATT – *treatment effect on the treated*)

$$\Delta_{ATT} = E(\Delta | D = 1) = E(Y^1 | D = 1) - E(Y^0 | D = 1). \quad (4)$$

Średni efekt interwencji jest różnicą pomiędzy oczekiwanym efektem dla uczestników interwencji oraz dla grupy osób nie poddanych działaniu instrumentu. Istotną wadą tej stosunkowo łatwej do oszacowania miary jest fakt, że dotyczy ona losowo wybranej osoby bez względu na to, czy była ona objęta programem, czy też nie. Badacze, oceniając konkretny instrument, muszą wziąć przede wszystkim pod uwagę efekt w grupie osób uczestniczących w danym projekcie. Poszukiwany jest zatem średni efekt interwencji dla osób poddanych interwencji, który stanowi różnicę pomiędzy średnim efektem dla osób objętych oddziaływaniem oraz średnim efektem dla osób objętych oddziaływaniem w sytuacji, gdyby te osoby nie były objęte oddziaływaniem.

Drugi z efektów jest nieobserwowalny i musi być imputowany z dostępnych danych. W praktyce można oszacować średni efekt zaobserwowany po zakończeniu interwencji dla osób, które w niej nie uczestniczyły ($E(Y^0 | D = 0)$). Gdyby zachodziła relacja (Caliendo, 2006):

$$E(Y^1 | D = 1) = E(Y^0 | D = 0), \quad (5)$$

osoby, które nie zostały poddane interwencji, byłyby traktowane jako grupa kontrolna dla uczestników programów. W metodach nieeksperymentalnych warunek ten nie jest spełniony, czyli:

$$E(Y^1 | D = 1) \neq E(Y^0 | D = 0). \quad (6)$$

Naturalną konsekwencją szacowania Δ_{ATT} jako różnicy średnich w subpopulacjach uczestników programu oraz osób nieuczestniczących jest obciążenie wynikające z selekcji (*selection bias*). Bezpośrednim źródłem tego obciążenia jest działanie mechanizmów selekcji, które odpowiadają za przypisanie jednostek do grupy interwencji, a ich konsekwencją jest występowanie różnic pomiędzy tymi jednostkami, które znalazły się w grupie badanej i grupie kontrolnej. Obciążenie wynika z faktu, że rozkłady cech uczestników interwencji oraz nieuczestników są różne. Cechy mogą mieć zarówno charakter obserwowalny (płeć, wiek, poziom wykształcenia itd.), jak i nieobserwowalny (poziom motywacji itp.). Cechy te z kolei mogą wpływać na partycypację danej jednostki w programie, a po jego zakończeniu na wynik. Efekt działania może zatem zależeć od rozkładu cech wpływających na podjęcie uczestnictwa (D) w formie interwencji i jednocześnie na jej efekt (Y)

w grupie osób poddanych i niepoddanych oddziaływaniu. Nieuwzględnienie mechanizmów selekcji podczas szacowania efektów podjętej interwencji bez korekt może spowodować błędną kalkulację efektu przyczynowego (Trzciniński, 2009).

Szczególnie efektywnym i szeroko wykorzystywanym narzędziem korekty obciążenia szacunków efektu jest, oparta na podejściu stanów kontryfaktycznych, technika PSM (*propensity score matching*).

1.1. TECHNIKA *PROPENSITY SCORE MATCHING*

Implementacja metody *propensity score matching* jest procedurą składającą się z kilku etapów. Rozpoczyna się od określenia populacji, dla której ma być szacowany efekt zastosowania instrumentu interwencji. Następnie odbywa się selekcja oraz pomiar zmiennych, które wpływają na uczestnictwo w programie oraz efekt tego programu po jego zakończeniu. Zakłada się, że proces selekcji nie jest losowym oraz odpowiada za mierzalne i niemierzalne różnice pomiędzy grupą uczestników i grupą osób, które nie uczestniczą w danym programie. Następnie spośród wszystkich uczestników programu dokonuje się losowego doboru jednostek do grupy badanej, a spośród wszystkich osób nieuczestniczących losowego doboru jednostek do znacznie większej grupy kontrolnej. Następnie na podstawie połączonych prób osób z grupy badanej i puli kontrolnej dokonuje się oszacowania *propensity scores*, które w istocie są prawdopodobieństwami uczestnictwa w określonej formie interwencji. W kroku kolejnym do każdej osoby z grupy badanej dołącza się za pomocą mechanizmów łączenia osobę lub osoby z puli kontrolnej. Po zakończeniu tego etapu ocenia się jakość dopasowania i dokonuje pomiaru efektu przyczynowego zastosowanej interwencji. Do krytycznych etapów zalicza się dobór i pomiar zmiennych wpływających na partycypację w programie oraz jego wynik, estymację skłonności do uczestnictwa w programie, selekcję grupy kontrolnej oraz ocenę jakości dopasowania grupy kontrolnej do grupy badanej po zaimplementowaniu mechanizmów selekcji (Konarski, Kotnarowski, 2007).

Celem metody PSM jest znalezienie dla każdej osoby z grupy badanej osoby jak najbardziej podobnej z puli kontrolnej ze względu na cechy reprezentowane przez zmienne X_1, X_2, \dots, X_q . Z uwagi na fakt, że nawet w dużym zbiorze osób nieuczestniczących w danym programie mogą nie wystąpić przypadki podobne, ze względu na cechy X_1, X_2, \dots, X_q , do osób poddanych programowi, rezygnuje się z dokładnego dopasowania na rzecz dopasowania osób, które są tylko w pewnym stopniu podobne do uczestników formy interwencji. To podobieństwo wyrażane jest za pomocą *propensity score*, która definiowana jest jako skłonność do uczestnictwa w programie i formalnie

wyrażana prawdopodobieństwem uczestnictwa w programie. Prawdopodobieństwo to jest szacowane na podstawie modelu, w którym zmiennymi objaśniającymi są cechy X_1, X_2, \dots, X_q , a zmienną objaśnianą D , przyjmująca wartość 1, gdy dana osoba uczestniczyła w programie. Najczęściej do oszacowania tych prawdopodobieństw wykorzystuje się model logitowy. *Propensity score* dla określonej osoby s jest prawdopodobieństwem, że osoba charakteryzowana wektorem wartości cech X_s będzie należeć do grupy uczestników programu:

$$P(X_s) = \Pr(D = 1 | X_s). \quad (7)$$

Propensity score dla konkretnej osoby jest skalarem, opisującym w sposób zsyntetyzowany za pomocą jednej wartości wiele (q) cech. W ten sposób problem doboru jednostek kontrolnych do osób badanych z zadania q -wymiarowego jest redukowany do zadania jednowymiarowego.

Uzyskując grupę kontrolną za pomocą techniki PSM, można oszacować efekt przyczynowy interwencji. Estymacji dokonuje się dla każdej osoby w grupie interwencji przez zestawienie efektu uczestnictwa w grupie interwencji z ważonym efektem osób dołączonych z puli kontrolnej. Efekt szacowany jest wg wzoru (Caliendo, 2006):

$$\Delta^{MAT} = \frac{1}{N_1} \sum_{i \in I_1} \left[Y_i^1 - \sum_{j \in I_0} w_{ij} Y_j^0 \right], \quad (8)$$

gdzie:

Δ^{MAT} jest efektem netto zastosowania interwencji,

N_1 jest liczebnością grupy poddanej interwencji,

I_1 jest zbiorem i uczestników interwencji,

I_0 jest zbiorem j osób nieuczestniczących w interwencji,

w_{ij} jest wagą przypisaną j -tej osobie z grupy kontrolnej, wykorzystywaną w wyznaczaniu kontrfaktycznego efektu dla i -tej osoby z grupy interwencji.

Dla każdej osoby z grupy interwencji suma wag w_{ij} osób przypisanych z grupy kontrolnej musi być równa 1, czyli:

$$\forall_i \sum_j w_{ij} = 1. \quad (9)$$

W celu wyznaczenia grupy kontrolnej dla każdej osoby z grupy poddanej interwencji definiuje się sąsiedztwo $C(P_i)$. Jednostki dołączone do i -tego uczestnika interwencji tworzą zbiór $A_i = \{j \in I_0 \mid P_j \in C(P_i)\}$. P_i oraz P_j oznaczają wartości *propensity score* odpowiednio dla i -tego uczestnika in-

terwencji oraz j -tej z grupy osób nieuczestniczących w interwencji. Różne podejścia do określania sąsiedztwa $C(P_i)$ oraz konstrukcji wag w_{ij} wyznaczają wachlarz możliwych sposobów określania grupy kontrolnej.

Najczęściej wykorzystywaną techniką wyznaczania grupy kontrolnej jest metoda najbliższego sąsiada (*nearest neighbor method*). W metodzie tej grupę kontrolną wyznaczają wartości (Heckmann, Ichimura, Todd, 1998):

$$C^{NN}(P_i) = \min_j \|P_i - P_j\|, j \in N_0, \quad (10)$$

gdzie:

$\|\dots\|$ jest określoną metryką,

N_0 jest liczebnością puli kontrolnej.

W podstawowej wersji tej metody najpierw porządkuje się losowo jednostki poddane interwencji, a następnie dla każdej osoby z grupy interwencji z wartością P_i wybiera się spośród nieuczestników osobę z najbardziej zbliżoną do P_i wartością P_j . Wagi wykorzystywane w estymacji efektu przyczynowego przyjmują wówczas wartość:

$$w_{ij}^{NN} = \begin{cases} 1; & \|P_i - P_j\| = \min_j \|P_i - P_j\| \\ 0; & \text{w przeciwnym przypadku} \end{cases} \quad (11)$$

Metoda najbliższego sąsiada w tym wariantie jest szczególnie efektywna, gdy pula kontrolna jest stosunkowo liczna, a więc wszystkie jednostki w grupie interwencji mają szansę znalezienia swoich bliskich odpowiedników.

Zaproponowano szereg modyfikacji metody najbliższych sąsiadów. Można ją stosować w wersji ze zwracaniem (*with replacement*), w której osoba wyselekcjonowana z puli kontrolnej (spośród wszystkich osób nieuczestniczących w interwencji) i przypisana do określonej osoby z grupy interwencji powraca do puli kontrolnej i może zostać ponownie dopasowana do innej osoby z grupy poddanej działaniu instrumentu. W wariantie bez zwracania (*without replacement*) raz dopasowana do osoby z grupy objętej interwencją osoba z puli kontrolnej zostaje z tej puli usunięta (Smith, Todd, 2005). Kolejna modyfikacja metody najbliższego sąsiada polega na doborze do każdej jednostki z grupy badanej więcej niż jednej osoby z puli kontrolnej, tzn. zamiast łączenia 1 do 1 (1:1) stosuje się łączenie 1 do wielu (1:k). Decydując się na zastosowanie tego podejścia, badacz musi określić liczbę osób k z puli kontrolnej, które muszą być dołączone do każdej jednostki i z grupy badanej. Ponadto musi zostać określona reguła ustalania wag w_{ij} . Jedną z możliwości daje zastosowanie wag jednostajnych (*uniform weights*),

które dla każdej jednostki dołączonej do i -tej jednostki z grupy interwencji (należącej do zbioru A_i) przyjmują stałą wartość $\frac{1}{k}$, pozostałe natomiast 0 (Davies, Kim, 2004):

$$w_{ij}^{NN} = \begin{cases} \frac{1}{k}; & \text{gdzie } j \in A_i \\ 0; & \text{w przeciwnym przypadku} \end{cases} \quad (12)$$

Alternatywnym rozwiązaniem względem metody najbliższego sąsiada jest podejście oparte na estymatorach jądrowych.

2. ESTYMATORY JĄDROWE W EWALUACJI AKTYWNYCH PROGRAMÓW RYNKU PRACY

W metodzie opartej na estymatorach jądrowych do oszacowania kontrfaktycznego wyniku danej osoby wykorzystywane są wyniki wszystkich osób, które należą do grupy kontrolnej ($A_i=I_0$), a nie jak to miało miejsce w technice *propensity score matching* tylko pewien podzbiór osób należących do puli kontrolnej. Znaczenie poszczególnych osób kontrolnych w kształtowaniu kontrfaktycznego wyniku zależy od bliskości do porównywanej osoby z grupy badanej. Wyższą wagę otrzymują osoby o zbliżonych do P_i wartościach P_j . W wariancie lokalnie stałym waga osób kontrolnych liczona jest wg formuły (Heckmann, Ichimura, Todd, 1997):

$$w_{ij}^{KM} = \frac{k\left(\frac{P_i - P_j}{h}\right)}{\sum_j k\left(\frac{P_i - P_j}{h}\right)}, \quad (13)$$

gdzie:

$k(\cdot)$ jest funkcją jądrową,

h jest parametrem wygładzania.

Funkcja jądrowa $k(y)$ musi posiadać cechę (Parzen, 1962):

$$\int k(y)dy = 1. \quad (14)$$

Zakłada się również, iż funkcja $k(y)$ jest symetryczna względem zera, czyli:

$$k(y) = k(-y) \quad (15)$$

i ma w tym punkcie słabe maksimum globalne:

$$k(0) \geq k(y). \quad (16)$$

Do konstrukcji estymatorów jądrowych używa się funkcji gęstości rozkładów zmiennych losowych. Najczęściej stosowanymi jądrami są:

– jądro rozkładu normalnego

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, \quad (17)$$

– jądro Cauchy'ego

$$k(z) = \frac{2}{\pi(z^2 + 1)^2}. \quad (18)$$

Optymalną wartość parametru wygładzania (inne nazwy to: szerokość pasma albo szerokość okna) h można wyznaczyć ze wzorów:

$$h = 1,059sn^{-1/5}, \quad (19)$$

$$h = 0,785(\hat{q}_3 - \hat{q}_1)n^{-1/5}, \quad (20)$$

gdzie:

s – odchylenie standardowe zmiennej P_i ,

$\hat{q}_3 - \hat{q}_1$ – rozstęp kwartyłowy zmiennej P_i .

Wartości h uzyskane ze wzorów (19) i (20) sprawdzają się w wielu przypadkach, jednakże gdy rozkład zmiennej jest bimodalny albo ma dużą skośność, mogą mieć tendencję do przeszacowania. W takich sytuacjach proponuje się używanie „reguły kciuka” (*thumb rule*) danej wzorem (Silverman, 1986):

$$h = 0,9 \min\left(s, \frac{\hat{q}_3 - \hat{q}_1}{1,349}\right)n^{-1/5}. \quad (21)$$

Bardziej złożonym wariantem metody opartej na estymatorach jądrowych jest podejście, w którym łączenia oraz oszacowania efektów netto interwencji dokonuje się za pomocą wag wyznaczonych estymatorem lokalnie liniowym (*local linear*) (Fan, 1993). Zasadniczą zaletą tego podejścia jest szybszy stopień zbieżności wokół punktów granicznych. Wagi w tym podejściu zdefiniowane są wzorem (Heckmann, Ichimura, Todd, 1998):

$$w_{ij}^{LL} = \frac{k_{ij} \sum_{k \in I_0} k_{ik} (P_k - P_i)^2 - k_{ij} (P_j - P_i) \sum_{k \in I_0} k_{ik} (P_k - P_i)}{\sum_{j \in I_0} k_{ij} \sum_{k \in I_0} k_{ij} (P_k - P_i)^2 - \left(\sum_{k \in I_0} k_{ik} (P_k - P_i) \right)^2}. \quad (22)$$

gdzie: $k_{ij} = k \left(\frac{P_i - P_j}{h} \right)$.

W istocie za pomocą estymatorów jądrowych dokonuje się oszacowania wyrazu wolnego w modelu regresji dla zmiennej Y_j^0 z jądrowymi wagami. Wagi zależą od odległości pomiędzy każdą jednostką z puli kontrolnej a uczestnikiem interwencji, dla którego wyznaczany jest stan kontrfaktyczny. Oszacowaniem średniego efektu dla stanów kontrfaktycznych jest wyraz wolny. Zasadnicza różnica pomiędzy estymatorem lokalnie stałym i lokalnie liniowym polega na tym, że w drugim z nich obok wyrazu wolnego uwzględnia się komponent liniowy ze względu na P_i (Caliendo, 2006).

Zarówno estymator lokalnie stały, jak i estymator lokalnie liniowy są szczególnymi przypadkami estymatora lokalnie wielomianowego (*local polynomial*) dowolnego rzędu p , który obok wyrazu wolnego zawiera również składnik wielomianowy stopnia p ze względu na P_i . Wagi lokalnie wielomianowe dowolnego stopnia p wygodnie jest przedstawić w zapisie macierzowym (Wand, Jones, 1995):

$$w_{ij}^{LP} = \mathbf{e}_1' (\mathbf{X}_i' \mathbf{\Omega}_i \mathbf{X}_i)^{-1} \mathbf{X}_i' \mathbf{\Omega}_i, \quad (23)$$

gdzie:

\mathbf{e}_1 – wektor jednostkowy o wymiarach $(p+1) \times 1$ z wartością 1 na pierwszej współrzędnej oraz 0 na pozostałych,

$\mathbf{\Omega}_i$ – diagonalna macierz wag jądrowych o wymiarach $(N_0 \times N_0)$ postaci:

$$\mathbf{\Omega}_i = \text{diag} \left\{ \frac{1}{h} k \left(\frac{P_1 - P_i}{h} \right), \dots, \frac{1}{h} k \left(\frac{P_{N_0} - P_i}{h} \right) \right\},$$

\mathbf{X}_i – macierz o wymiarach $(N_0 \times (p+1))$ postaci:

$$\mathbf{X}_i = \begin{bmatrix} 1 & P_1 - P_i & \dots & (P_1 - P_i)^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & P_{N_0} - P_i & \dots & (P_{N_0} - P_i)^p \end{bmatrix}.$$

3. WYNIKI ANALIZ EMPIRYCZNYCH

Analizę empiryczną przeprowadzono na danych o bezrobotnych zarejestrowanych w Powiatowym Urzędzie Pracy w Toruniu i uczestniczących w aktywnych programach rynku pracy w 2009 oraz 2010 r. Analizę metodą *propensity score matching* przeprowadzono dla wariantu łączenia 1:1 bez zwracania i ze zwracaniem. W analizie opartej na estymatorach jądrowych zastosowano jądro Gaussa oraz Cauchy'ego. Parametr wygładzania wyznaczono za pomocą formuły danej wzorem (21).

Tabela 1. Oceny efektywności netto aktywnych programów rynku pracy dokonane techniką *propensity score matching*

rok	aktywny program rynku pracy	liczba uczestników programu	liczba zatrudnionych po programie	efekt brutto	efekt netto	
					bez zwracania	ze zwracaniem
2009	Szkolenia	385	205	0,5325	0,0519	0,0721
	Staże	628	433	0,6895	0,2150	0,2123
	Prace interwencyjne	26	12	0,4615	-0,1923	-0,0501
	Prace społecznie użyteczne	115	20	0,1739	-0,0870	-0,0480
	Roboty publiczne	20	5	0,2500	-0,2000	-0,0868
	Dotacje	154	154	1,0000	0,6039	0,5929
	Wszystkie	1328	829	0,6242	0,1438	0,1718
2010	Szkolenia	677	296	0,4372	-0,0074	0,0106
	Staże	950	589	0,6200	0,1832	0,1858
	Prace interwencyjne	42	23	0,5476	0,2381	0,1652
	Prace społecznie użyteczne	123	16	0,1301	-0,0894	-0,0975
	Roboty publiczne	24	12	0,5000	0,2083	0,2467
	Dotacje	213	213	1,0000	0,5634	0,564
	Wszystkie	2029	1149	0,5663	0,1419	0,1562

Źródło: obliczenia własne.

W 2009 r. w aktywnych programach rynku pracy uczestniczyło 1328 osób, a w 2010 r. 2029 osób. Liczba osób, które znalazły zatrudnienie po uczestnictwie w programach aktywnych w 2009 r., wynosiła 829 osób. W 2010 r. nastąpił wzrost zatrudnionych do 1149 osób. Wzrostowi liczby uczestników w 2010 r. w stosunku do roku poprzedniego (o 52,8%) towarzyszył wzrost liczby zatrudnionych o 38,6%. Nieproporcjonalnie mniejszy wzrost liczby zatrudnionych po uczestnictwie w programach aktywnych

w stosunku do wzrostu liczby uczestników spowodował, że wskaźniki efektywności programów rynku pracy zmalały w 2010 r. w stosunku do 2009 r.

Biorąc pod uwagę efekty brutto zarówno w 2009, jak i w 2010 r., należy zaznaczyć, że najmniej efektywne były prace społecznie użyteczne. Pełen sukces dały dotacje na podjęcie działalności gospodarczej, co jest spowodowane tym, że podjęcie działalności gospodarczej jest równoznaczne z podjęciem zatrudnienia.

Szacunki efektywności netto szkoleń uzyskane za pomocą estymatorów jądrowych są zbliżone z wynikiem uzyskanym techniką PSM w wariancie ze zwracaniem.

Szacunki efektywności netto staży dla wszystkich metod są podobne i oscylują wokół wartości 0,2 dla 2009 r. Dla 2010 r. przyjmują wartości nieznacznie mniejsze, przy czym szacunki uzyskane za pomocą estymatorów jądrowych są większe od tych uzyskanych metodą PSM.

Tabela 2. Oceny efektywności netto aktywnych programów rynku pracy dokonane za pomocą estymatorów jądrowych

rok	aktywny program rynku pracy	jądro Gaussa			jądro Cauchy'ego		
		lokalnie stały	lokalnie liniowy	lokalnie kwadratowy	lokalnie stały	lokalnie liniowy	lokalnie kwadratowy
2009	Szkolenie	0,0734	0,0744	0,0721	0,0819	0,0744	0,0731
	Staż	0,1993	0,2032	0,1995	0,2126	0,2000	0,2036
	Prace interwencyjne	-0,0251	-0,0362	-0,0466	-0,0053	-0,0357	-0,0464
	Prace społecznie użyteczne	-0,0913	-0,0556	-0,0378	-0,1768	-0,0546	-0,0433
	Roboty publiczne	-0,1412	-0,1143	-0,1068	-0,1626	-0,1016	-0,1018
	Dotacje	0,5645	0,5664	0,563	0,5648	0,5675	0,5642
	Wszystkie	0,1817	0,1813	0,1825	0,1876	0,1794	0,1807
	2010	Szkolenie	0,0130	0,0123	0,0122	0,018	0,0123
Staż		0,1918	0,1937	0,1908	0,1979	0,1919	0,1944
Prace interwencyjne		0,1455	0,1318	0,1454	0,1574	0,1323	0,1413
Prace społecznie użyteczne		-0,1018	-0,076	-0,0626	-0,1864	-0,0750	-0,0673
Roboty publiczne		0,1621	0,1797	0,1858	0,1352	0,1799	0,1839
Dotacje		0,5725	0,5719	0,5679	0,582	0,5728	0,5683
Wszystkie		0,1609	0,1645	0,1613	0,164	0,1615	0,1619

Źródło: obliczenia własne.

Ocena efektywności prac interwencyjnych wskazuje, że w 2009 r. był to instrument nieefektywny. Zastosowanie każdej metody dało ujemne estyma-

tory efektu netto. Dla 2010 r. natomiast oceny efektywności są dodatnie i przyjmują dość duże wartości. Szacunki uzyskane za pomocą estymatorów jądrowych są nieznacznie większe od wyników uzyskanych techniką PSM.

W przypadku prac społecznie użytecznych w obu analizowanych okresach wystąpiły ujemne efekty zatrudnieniowe, co oznacza, że częściej znajdowały zatrudnienie osoby, które nie uczestniczyły w tym instrumencie aktywnej polityki rynku pracy. Szacunki uzyskane za pomocą estymatora jądrowego lokalnie stałego przyjmują znacznie niższe wartości w porównaniu z wynikami uzyskanymi estymatorami lokalnie liniowym i lokalnie kwadratowym.

Roboty publiczne, podobnie jak prace interwencyjne, były instrumentem nieefektywnym w 2009 r. W 2010 r. natomiast dały silne efekty zatrudnieniowe. Szacunki efektywności uzyskane estymatorami jądrowymi dla 2009 r. przyjęły wartości znacznie wyższe od tych wyznaczonych techniką PSM w wariancie bez zwracania oraz niższe od uzyskanych techniką PSM ze zwracaniem. Dla 2010 r. wszystkie oceny efektywności uzyskane estymatorami jądrowymi są niższe niż uzyskane techniką PSM.

Efekty netto dotacji na podjęcie działalności gospodarczej w każdym przypadku przyjmują wysokie wartości przekraczające 0,55. Dla 2009 r. estymatory jądrowe dały nieznacznie niższe oceny efektów netto aniżeli technika PSM. W 2010 r. wszystkie oceny są ze sobą zbieżne.

Efekty zatrudnieniowe wszystkich analizowanych programów łącznie były relatywnie duże. W 2010 r. efekty te były niższe w porównaniu z efektami dla 2009 r. Szacunki efektywności uzyskane za pomocą estymatorów jądrowych w każdym przypadku dały wyższe wartości w porównaniu z ocenami osiągniętymi techniką *propensity score matching*.

PODSUMOWANIE

Na podstawie przeprowadzonych analiz nie można wskazać przewagi którejkolwiek z przedstawionych metod szacowania efektywności netto aktywnych programów rynku pracy nad pozostałymi. Zarówno technika *propensity score matching*, jak i podejście oparte na estymatorach jądrowych dały podobne wyniki. Nieznaczne różnice wynikają z przyjętych metod analizy.

LITERATURA

- Bieliński J., Bober M., Sarzalska M., Zawistowski J. (2008), *Aktywne polityki na elastycznym rynku pracy*, [w:] Bukowski M. (red.), *Zatrudnienie w Polsce w 2007. Bezpieczeństwo na elastycznym rynku pracy*, Departament Analiz Ekonomicznych i Prognoz, Ministerstwo Pracy i Polityki Społecznej, Warszawa 2008.
- Caliendo M. (2006), *Microeconomic Evaluation of Labour Market Policies*, Springer, Berlin.
- Davies R., Kim S. (2004), *Matching and the Estimated Impact of Interlisting*, ISMA Working Paper.
- Fan J. (1993), *Local Linear Regression Smoothers and Their Minimax Efficiencies*, „The Annals of Statistics”, 21(1), 196–216,
DOI: <http://dx.doi.org/10.1080/01621459.1994.10476781>.
- Heckman J., Ichimura H., Todd P. (1998), *Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program*, „Review of Economic Studies”, 65(2), 261–294, DOI: <http://dx.doi.org/10.1111/1467-937X.00044>.
- Konarski R., Kotnarowski M. (2007), *Zastosowanie metody propensity score matching w ewaluacji ex-post*, [w:] Haber A. (red.), *Ewaluacja ex-post. Teoria i praktyka badawcza*, PARP, Warszawa.
- Maksim M. (2008), *Ewaluacja szkoleń dla bezrobotnych w województwie kujawsko-pomorskim*, TNOiK, Toruń.
- Parzen E. (1962), *On Estimation of a Probability Density Function and Mode*, „The Annals of Mathematical Statistics”, 33(3), 1065–1076.
- Silverman B. W. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.
- Smith J. A., Todd P. E. (2005), *Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?*, „Journal of Econometrics”, 125, 305–353,
DOI: <http://dx.doi.org/10.2139/ssrn.286297>.
- Trzeciński R. (2009), *Wykorzystanie techniki propensity score matching w badaniach ewaluacyjnych*, Polska Agencja Rozwoju Przedsiębiorczości, Warszawa.
- Wand M. P., M. C. Jones (1995), *Kernel smoothing*, Chapman and Hall, London.

APPLICATION OF KERNEL ESTIMATORS TO ESTIMATION
EFFICIENCY OF ACTIVE LABOR MARKET PROGRAMS

A b s t r a c t. Each intervention carried out by the state of the labor market must be evaluated at the end in terms of efficiency. The purpose of this article is to evaluate the effectiveness of active labor market policies. As a research tool will be used Propensity Score Matching technique and an approach based on kernel estimators. Theoretical considerations are supported by empirical analysis.

K e y w o r d s: net effectiveness of labor market programs, kernel estimator.