

# Empiryczna weryfikacja założeń metodologii ewaluacji wpływu do potrzeb oceny inwestycji współfinansowanych z funduszy unijnych

*Dominika Wojtowicz, Michał Wolański*

*Celem artykułu jest przedstawienie założeń ewaluacji wpływu oraz metodologii szacowania efektu netto interwencji publicznych jako narzędzia podwyższającego jakość oceny projektów realizowanych przy współudziale środków unijnych. Prezentację wyników empirycznych badań dotyczących zastosowania tej metodologii przy ocenie wyników inwestycji w sektorze infrastruktury transportowej poprzedza analiza dotychczasowego zastosowania tej metodologii.*

## 1. Wstęp

Od momentu przystąpienia do Unii Europejskiej w 2004 r. Polska stała się podmiotem wspólnej polityki spójności, mającej na celu dążenie do osiągnięcia kohezji w trzech wymiarach – gospodarczym, społecznym i przestrzennym. Na poziomie poszczególnych państw przy wsparciu funduszy unijnych realizowane są programy i projekty wpisujące się założenia dokumentów strategicznych – w Polsce Narodowej Strategii Spójności. W latach 2007–2013 Polska jest największym beneficjentem polityki spójności, otrzymując łącznie 67,3 mld EUR, czyli 19,4% całego budżetu przeznaczanego na tę politykę, wynoszącego 347,4 mld EUR.

W związku z tym niezmiernie istotne jest dołożenie wszelkich starań, aby środki przekazywane Polsce były jak najtrafniej alokowane – wspierały interwencje przynoszące jak najwięcej pozytywnych efektów. Ocena tych efektów jest dokonywana w procesie ewaluacji, a więc szacowania wartości realizowanych programów i projektów. Według K. Olejniczaka ewaluację można zdefiniować jako zróżnicowany zestaw metod i działań zorientowanych na krytyczną refleksję nad wartością i jakością interwencji publicznych – zarówno procesów ich wdrażania, jak też ich efektów (Olejniczak, Kozak i Ledzion 2008: 19).

W niniejszym artykule autorzy podejmują problematykę związaną z głównymi założeniami oraz możliwościami praktycznego zastosowania metodologii tzw. ewaluacji wpływu (*impact evaluation*) przy ocenie projektów reali-

zowanych przy współudziale środków unijnych. Rozważania o charakterze teoretycznym poparte są wynikami badań zrealizowanych w 2008 r. w ramach projektu naukowego, który miał na celu wzmocnienie potencjału ewaluacyjnego Polski w zakresie oszacowywania efektów netto programów i projektów współfinansowanych z funduszy unijnych. Badania koncentrowały się na możliwościach adaptacji metodologii pomiaru efektu netto interwencji publicznych w infrastrukturę transportu drogowego.

Proponowana metodologia nie była dotychczas stosowana w sektorze infrastrukturalnym ze względu na szereg problemów związanych z trudnościami z wyznaczeniem wiarygodnych współczynników sukcesu (np. w przypadku budowy skrzyżowań), zebraniem danych czy wreszcie wyborem tzw. sytuacji kontrfaktycznych (*counterfactual situations*). Przez to ostatnie pojęcie rozumie się jednostki lub populacje możliwie podobne do danej, które nie zostały poddane interwencji. Wzajemne ich porównanie pozwala określić zmiany, niebędące wynikiem ewaluowanych działań.

## 2. Główne założenia metodologii szacowania efektu netto

Jednym z kluczowych zadań ewaluacji jest ustalenie związku przyczynowego między podjętą interwencją a obserwowanymi wynikami. Ustalenie przyczynowego wpływu interwencji wymaga (Górniak 2007):

- zmierzenia efektu całkowitego interwencji jako ogólnej zmiany w zakresie zdefiniowanym przez cel (zmiany na poziomie przyjętych wskaźników),
- oddzielenia zmiany niezależnej od podjętej interwencji od zmiany, którą można przypisać wpływowi interwencji (efekt netto).

*Impact evaluation* opierająca się na pomiarze efektu netto jest jednym z bardziej zaawansowanych narzędzi ewaluacji skutków inwestycji publicznych, gdyż pomaga ocenić jej rzeczywisty efekt, oddzielając oddziaływanie czynników zewnętrznych. Niemniej nawet zwolennicy i propagatorzy tej metody (w głównej mierze reprezentanci włoskich ośrodków naukowych) podkreślają, że jej skuteczne zastosowanie ma wiele ograniczeń i w praktyce szacowanie efektów netto najlepiej sprawdza się przy ewaluacji projektów z zakresu podnoszenia jakości zasobów ludzkich i przedsiębiorczości (Retore, Trivellato i Martini 2003: 302).

Szacowanie efektu netto jest oparte na metodologii zbliżonej do eksperymentu. Istotą eksperymentu jest zbadanie wpływu zmiennej niezależnej (np. przebudowy skrzyżowania) na zmienną zależną (np. organizacja ruchu na skrzyżowaniu i jego funkcjonowanie). Z reguły w badaniach eksperymentalnych zmienna niezależna przybiera formę bodźca eksperymentalnego, który albo występuje, albo nie. Oznacza to, że jest ona zmienną dychotomiczną, przyjmującą dwie możliwe wartości: obecna lub nieobecna. W typowym schemacie eksperymentator porównuje sytuację, w której oddziałuje

bodźcem z sytuacją, w której bodźca brak (Babbie 2007: 247). W przypadku interwencji publicznych zastosowanie typowego eksperymentu byłoby niemożliwe, niemniej oparcie koncepcji badania na metodzie eksperymentu jest w pełni uzasadnione i często praktykowane. Tak jak w przypadku eksperymentu na podstawie wybranych kryteriów badane jest funkcjonowanie podmiotu kontrfaktycznego, a więc takiego, którego parametry są możliwie najbardziej zbliżone do parametrów podmiotu poddanego interwencji publicznej.

Należy podkreślić, że podejście metodologiczne *impact evaluation* jest zgodne ze stanowiskiem Komisji Europejskiej, która podkreśla, że „Gdy mówimy, że dane efekty zostały wyprodukowane lub spowodowane przez program, oznacza to, że gdyby program nie miał miejsca lub miałby miejsce w innej formie lub innym stopniu, efekty te nie wystąpiłyby lub nie wystąpiłyby w tym samym stopniu. Oznacza to, że ważne jest, by mieć precyzyjny obraz tego, co stałoby się bez programu” (Komisja Europejska 1997: 39).

### 3. Zastosowanie metodologii szacowania efektu netto w dotychczasowych ewaluacjach

Pomimo zaznaczenia przez Komisję Europejską w *Przewodniku ewaluacji w połowie okresu i ex-post* (Komisja Europejska 1997: 21) zalet stosowania metody analizy sytuacji kontrfaktycznych pozwalających na badanie efektu netto programów lub projektów w ramach badań ewaluacyjnych, w rzeczywistości narzędzie to wykorzystywane jest w bardzo ograniczonym zakresie, czego dowodzi przeprowadzone przez autorów niniejszego raportu analiza raportów *ex-post* i *mid term* programów realizowanych w wybranych krajach UE (Wielka Brytania, Niemcy, Szwecja, Włochy) oraz liczne publikacje naukowe (Center for Global Development 2006; White 2006).

Większość badań dotyczących oceny efektów działań publicznych (polityk, programów, przedsięwzięć itp.) z wykorzystaniem metod porównania stanu rzeczywistego ze stanem kontrfaktycznym została zleconych przez Bank Światowy. Instytucja ta jest podmiotem bardzo aktywnie wspierającym badania nad zasadnością i korzyściami płynącymi ze stosowania *impact evaluation*, a więc ewaluacji pozwalającej na określenie realnych zmian w życiu beneficjentów programu/projektu. Ewaluacja wpływu pozwala odpowiedzieć na pytanie, jak wyglądałaby sytuacja osób poddanych interwencji, gdyby w rzeczywistości nie skorzystali z niej. Ogromnemu doświadczeniu zdobytemu przez Bank Światowy w zakresie prowadzenia tego typu badań dowodzi obszerna baza artykułów, raportów, studia przypadków i różnego rodzaju publikacji zamieszczonych na oficjalnej stronie instytucji (<http://web.worldbank.org>). Wśród badań przeprowadzonych w oparciu o zasady ewaluacji wpływu zleconych przez Bank Światowy dominują te, które dotyczą interwencji z zakresu polityki społecznej, polityk i działań

podejmowanych na rzecz rozwiązywania problemów biedoty w krajach trzeciego świata.

Niemniej znajdują się tam również dwa ciekawe opracowania zawierające wskazówki do koncepcji szacowania efektów netto projektów dotyczących dróg wiejskich (Walle 2009) oraz transportu miejskiego (Boarnet 2006). W pierwszej pracy, dotyczącej wpływu dróg wiejskich, autorka proponuje analizę, np. dochodów gospodarstw domowych lub wysokości ich konsumpcji, oszczędności czasu podróży, oszczędności kosztów podróży, zmiany cen na rynku lokalnym, mobilności mieszkańców, kierunków i wielkości ruchu migracyjnego. W drugim opracowaniu, dotyczącym transportu miejskiego, autor proponuje analizę różnych pól wpływu, takich jak: rynek pracy, korzystanie z usług oraz ceny ziemi. W opracowaniach tych pokazany jest również szereg metod możliwych do zastosowania przy ocenie efektów netto.

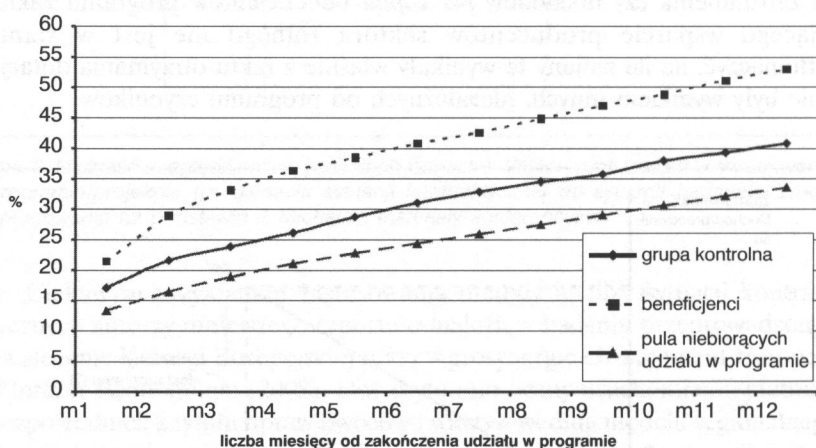
Ciekawą bazę raportów i publikacji z zakresu ewaluacji wpływu zawiera także strona internetowa Governance and Social Development Resource Centre (<http://www.gsdr.org>). Założona przez Department for International Development (Departament ds. Rozwoju Międzynarodowego) Wielkiej Brytanii jednostka zaleca stosowanie metod analizy stanów kontrfaktycznych jako narzędzia pozwalającego na określenie rzeczywistych efektów realizowanych programów i projektów na rzecz rozwoju państw oraz regionów zacofanych. Podkreśla przy tym konieczność prowadzenia tego typu ewaluacji ze względu na znaczne środki pochodzenia publicznego przeznaczane na interwencje z tego zakresu.

Inną jednostką promującą stosowanie metod analizy sytuacji kontrfaktycznych w badaniu efektów stosowanych przez rząd różnego rodzaju rozwiązań w sferze publicznej jest Jednostka Strategiczna (Strategic Unit) działająca przy premierze Wielkiej Brytanii. Na stronie internetowej instytucji (<http://interactive.cabinetoffice.gov.uk>), oprócz opisu głównych założeń metodologii badań ewaluacyjnych opartych na porównaniu sytuacji kontrfaktycznych dających możliwość szacowania realnego efektu netto danej interwencji, podane są przykłady ich zastosowania. Pierwszy z nich dotyczy oceny efektu wdrożenia dyrektywy unijnej dotyczącej składowania odpadów w Wielkiej Brytanii, drugi natomiast dowodzi skuteczności projektu „Restart interview for Jobseekers in 1989/90”, mającego na celu wsparcie bezrobotnych w powrocie na rynek pracy (grupa poddana interwencji 8000 osób, grupa kontrfaktyczna 500 osób).

Zastosowanie analizowanej metody jest również wykorzystywane przy ewaluacji interwencji publicznych w Stanach Zjednoczonych. Niemniej w większości tych badań analiza porównawcza nie wykorzystywała metodologii opartej na założeniach eksperymentu (porównanie stanów kontrfaktycznych), lecz koncentrowała się na ocenie efektywności danej interwencji w odniesieniu do podobnych projektów federalnych (dobór ze względu na ich przedmiot, zakres finansowy itp.) (<http://www.whitehouse.gov>).

Analiza opublikowanych w Polsce w ostatnich latach raportów ewaluacyjnych wskazuje, że zastosowanie metody szacowania efektu netto dotyczyło nieznaczącej części przeprowadzonych badań.

Metodę szacowania efektu netto na podstawie analizy stanów kontrfaktycznych wykorzystano w badaniu przeprowadzonym przez PBS na zlecenie Polskiej Agencji Rozwoju Przedsiębiorczości w kontekście oceny wpływu programów aktywizacji zawodowej. Efektem brutto był określony procent beneficjentów, którzy znaleźli pracę po określonym czasie od rozpoczęcia programu; za efekt netto uznano beneficjentów, którzy znaleźli pracę po określonym czasie od rozpoczęcia programu, a nie zaleźliby jej bez udziału w programie.



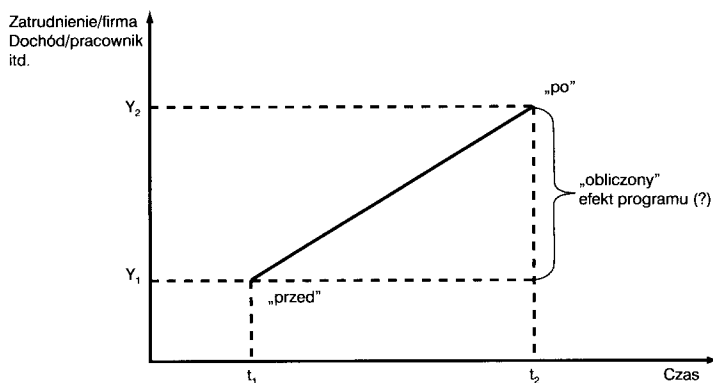
Rys. 1. Odsetek zatrudnionych po określonym czasie od udziału w programie aktywizacyjnym. Źródło: R. Konarski i M. Kontarowski 2007. Zastosowanie metody propensity score matching w ewaluacji ex-post, w: A Haber (red.) Ewaluacja ex-post. Teoria i praktyka badawcza, Warszawa: PARP.

Autorzy raportu dowiedli, że przy porównaniu odsetka zatrudnionych wśród beneficjentów i w grupie kontrolnej można zaobserwować, iż już w pierwszym miesiącu od zakończenia udziału w programie beneficjenci podejmowali pracę statystycznie istotnie częściej niż osoby należące do grupy kontrolnej. Porównanie zjawiska dla grupy biorącej udział w programie i grupy kontrolnej przeprowadzono dla kolejnych 12 miesięcy, co dało podstawy do wyciągnięcia rzetelnych wniosków dotyczących wystąpienia efektu netto realizowanego programu.

Innym przykładem zastosowania metody szacowania efektu netto jest ewaluacja typu ex-post Programu SAPARD w Polsce i na Słowacji prowadzona przez zespół z Uniwersytetu w Kiel (Niemcy) (Michalek i Zarnekow

2008). Należy zaznaczyć, że badacze zajmują się szerszym wykorzystaniem metody szacowania efektu netto przy użyciu analizy stanów kontrfaktycznych interwencji w sektorze rolnym (Michalek 2007; Michalek, Zarnekow, Stulrajter i Klank 2008).

Autorzy w raporcie metodologicznym bardzo wyraźnie podkreślają istotność zastosowania metody porównania stanów kontrfaktycznych jako jedynego narzędzia gwarantującego rzetelne oszacowanie rzeczywistego efektu netto danego programu/projektu. Najczęściej wykorzystywaną w badaniach ewaluacyjnych metodą badania efektów polegającą na zmierzeniu wybranych wskaźników przed i po realizacji danej interwencji nazywają „oceną naiwną” (*naive evaluation*), gdyż nie jest ona w stanie uwzględnić innych czynników (np. koniunktury, zmian w przepisach prawnych) oddziałujących na zjawisko/beneficjentów, których dotyczy projekt/program. I tak, zbadanie wzrostu zatrudnienia czy dochodów *per capita* beneficjentów programu zakładającego wsparcie producentów sektora rolnego nie jest w stanie wytłumaczyć, na ile zmiany te wynikały właśnie z faktu otrzymania dotacji, a nie były wynikiem innych, niezależnych od programu czynników.

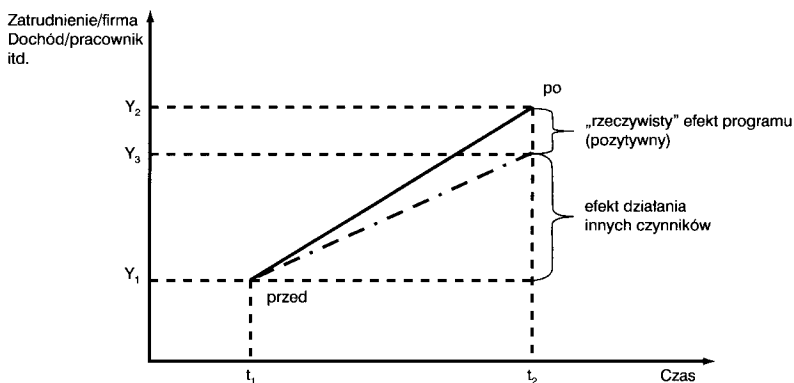


Rys. 2. Przykład „naiwnej” ewaluacji efektów programu. Źródło: opracowanie na podstawie J. Michalek, materiały konferencyjne.

Dlatego badacze powinni dokonać doboru grupy porównawczej. Podobieństwo jednostek poddanych interwencji i tych, które nie otrzymały wsparcia, powinno być zweryfikowane statystycznie na podstawie analizy rozkładu różnych obserwowanych zmiennych współzmiennych.

Zastosowanie metody analizy sytuacji kontrfaktycznych pozwala na udzielenie odpowiedzi na najważniejsze z punktu widzenia oceny rzeczywistych efektów projektu pytanie: „Jaka byłaby sytuacja beneficjentów, gdyby nie mieli oni możliwości skorzystania ze wsparcia w ramach programu?” (Micha-

lek, Zarnekow, Stulrajter i Klank 2008). Dzięki niej możemy wskazać na wystąpienie pozytywnego lub negatywnego efektu netto danej interwencji, co pokazuje rysunek 3.



Rys. 3. Istotność uwzględnienia w badaniach jednostek niebiorących udziału w ewaluowanym programie/projekcie (porównania sytuacji kontrafaktycznej do sytuacji badanej). Źródło: opracowanie na podstawie J. Michalek, materiały konferencyjne.

Ciekawym przykładem zastosowania metody analizy sytuacji kontrafaktycznych autorzy niniejszego raportu odnaleźli w badaniu przeprowadzonym na zlecenie Komisji Europejskiej przez Agrosynergie Groupement Européen d'Intérêt Economique (2008). Dotyczyło ono oceny uruchomienia płatności bezpośrednich z tytułu upraw owoców i warzyw według modelu regionalnego. Podejście ewaluacyjne opierało się na wynikach etapu strukturyzacji metody wynikającej z podstaw ekonomicznych. Taka analiza umożliwiła sformułowanie założeń, które zostały skonfrontowane z wynikami analizy ilościowej oraz jakościowej i pozwoliły na udzielenie odpowiedzi na pytania ewaluacyjne. Teoretyczne podejście mikroekonomiczne polegało na analizie zachowania producentów w odpowiedzi na różne sygnały z rynku i środki wsparcia.

Sytuacja kontrafaktyczna odpowiadająca sytuacji rynkowej bez środka wsparcia została porównana z trzema sytuacjami:

- całkowitym oddzieleniem płatności od produkcji,
- wyjątkami w stosowaniu całkowitego oddzielenia płatności od produkcji w postaci zasad użytkowania gruntów,
- wyjątkami w całkowitym oddzieleniu płatności od produkcji w postaci pomocy częściowo związanej z produkcją lub częściowo oddzielonej od produkcji.

Metodologia zastosowana do oceny była w tym badaniu połączeniem analizy ilościowej (klasyczna statystyczna analiza ilościowa i analiza modelizacyjna – pozytywne programowanie matematyczne) z gromadzeniem

i analizą informacji jakościowych. Analiza danych dotyczących różnych kierunków techniczno-ekonomicznych gospodarstw rolnych została dokonana dla sześciu regionów Unii Europejskiej (Emilia-Romania, Andaluzja, Akwitania, Wschodnia Anglia, Dania i Nadrenia-Północna Westfalia).

Dla uzyskania pełniejszego obrazu wykorzystania metody szacowania efektów netto interwencji publicznych autorzy dokonali przeglądu metodologii stosowanych w badaniach ewaluacyjnych wykonanych dotychczas w Polsce. Przeanalizowano wszystkie raporty typu *ex-post* znajdujące się w bazie badań ewaluacyjnych (<http://www.fundusze-strukturalne.gov.pl/NR/rdonlyres/8292BC86-7BE0-4D50-8F02-DEBECDC56AF/49191/bazabadan171008.xls> – stan na 1.10.2008). Na 12 raportów w 1 zastosowano analizę sytuacji kontrfaktycznych – w badaniu przeprowadzonym na zlecenie PARP w latach 2006–2007 „Ewaluacja *ex-post* komponentu regionalnego programu Phare 2002 SSG – Komponent RZL” porównano odsetek zatrudnionych w grupie beneficjentów, w grupie kontrolnej oraz w puli niebiorącej udziału w programie.

## 4. Szacowanie efektu netto inwestycji infrastrukturalnych – wyniki badań empirycznych

Ocena efektu netto w sektorze infrastruktury została zrealizowana na przykładzie projektu przebudowy skrzyżowań w całym kraju (finansowanego ze SPOT, zrealizowanego przez GDDKiA). Projekt w całości został zakończony w 2007 r. (ale niektóre skrzyżowania przed 2007 r.). Miał wartość 75 mln PLN, z czego 44 mln PLN pochodziły z środków UE. Dokonano wyboru tego projektu ze względu na jasny współczynnik sukcesu (zmniejszenie wypadkowości – funkcja wypadków, kolizji oraz ofiar).

Pomiar efektu netto badacze zrealizowali w kontekście efektu netto zmian wypadkowości – dla 24 skrzyżowań (przy czym liczba kolizji była rozpatrywana z podziałem na skrzyżowania, ale łączna wypadkowość została zmierzona ogólnie ze względu na małą liczbę wypadków śmiertelnych).

### 4.1. Dobór jednostek referencyjnych

Zazwyczaj do badania efektu netto w ewaluacjach korzysta się z metody *Propensity Score Matching* (PSM) (Konarski i Kontarowski 2007; Strawiński 2008). PSM to metoda, którą w latach 80. XX w. stworzono w celu wsparcia wyników badań eksperymentalnych. Chodziło m.in. o to, aby zniwelować efekt wpływu nielosowego doboru do grup eksperymentalnej i kontrolnej. W założeniu powinny być one dobrane losowo z określonej populacji (czasem bardzo dużej!), w praktyce jest to jednak niezbyt realne. W konsekwencji pojawia się trudny do oszacowania błąd, który może być na tyle duży, aby mieć znaczący wpływ na całkowity wynik badania. PSM zakłada wykorzystanie dodatkowych zmiennych towarzyszących i za ich pomocą wyrównanie różnic i błędu związanego z nielosowym doбором. Na ich pod-



stawie wylicza się metodą regresji logistycznej *propensity score* i dopasowuje obie grupy.

W przypadku niniejszego badania zachodziła uzasadniona obawa, że PSM nie da rzetelnych wyników bądź w ogóle nie da się zastosować ze względu na słabą jakość dopasowania. Obawa ta była oparta na przesłance dotyczącej faktu, że baza danych jednostek niepoddanych interwencji powinna być dość obszerna tak, by można było dobrać do jednostek ewaluowanych odpowiedniej jakości obserwacje bliźniacze.

Obawy co do rzetelności zastosowanych metod budziła także jakość bazy danych, jak również konieczność nielosowego doboru skrzyżowań referencyjnych. Dobór nielosowy był konieczny, gdyż skrzyżowania wybrane do przebudowy zostały określone, jako „szczególnie niebezpieczne”. Siłą rzeczy członkowie zespołu projektowego zmuszeni zostali do wyszukania również „niebezpiecznych” skrzyżowań referencyjnych. Stało się więc jasne, że nie można było wziąć pod uwagę całej populacji skrzyżowań, a jedynie wybrać spośród nich te, które – zdaniem badaczy – są szczególnie niebezpieczne. Konsekwencją takiego podejścia był celowy dobór jednostek do próby.

Zdecydowano się na klasyczny sposób doboru skrzyżowań referencyjnych, czyli na model regresji logistycznej. W modelu uwzględnione zostały następujące zmienne: liczba wlotów skrzyżowania, ruch na wszystkich wlotach skrzyżowania, rodzaj terenu (zabudowany, niezabudowany), rodzaj skrzyżowania (jednopoziomowe, dwupoziomowe, rondo) oraz liczba uszkodzonych w latach 2001–2005 (wypadki ogółem, zabici, ranni). Ostatecznie otrzymano dopasowanie skrzyżowań referencyjnych do wszystkich skrzyżowań ewaluowanych na poziomie 98,1%.

Na podstawie rozkładu prawdopodobieństw sukcesu zmiennej zależnej oszacowano bliskości poszczególnych skrzyżowań i ostateczny dobór skrzyżowań referencyjnych.

## **4.2. Szacowanie efektu netto inwestycji – analiza zgromadzonych danych**

Analizę danych przeprowadzano w sposób stopniowy. Najpierw wyliczono średnie roczne wartości wypadkowości w latach 2001–2005 dla skrzyżowań referencyjnych i ewaluowanych. W przypadku skrzyżowań, które wzięły udział w projekcie, był to okres przed przebudową. Następnie wyliczono wypadkowość w roku 2007 (czyli już po przebudowie 24 skrzyżowań, które wzięły udział w projekcie) i oszacowano proporcje mówiące o tym, jaki odsetek wypadkowości 2001–2005 stanowi wypadkowość z roku 2007. Otrzymane proporcje od siebie odjęto, otrzymując cztery składowe całościowego efektu netto. Były to: efekt netto dotyczący wypadkowości, liczby rannych oraz liczby zabitych.

Na ostatnim etapie podjęto próbę zredukowania efektu netto do dwóch tylko wskaźników: efektu netto liczby wypadków i efektu netto ich najważniejszych skutków, a zatem liczby uszkodzonych w wypadkach.

Zgodnie z zaproponowaną metodologią, dla 24 ewaluowanych skrzyżowań oraz dla skrzyżowań referencyjnych otrzymano wyniki zamieszczone w tabeli 1.

<b>Wypadki*</b>			
<b>Wyniki</b>	<b>średnia wypadków rocznie</b>	<b>średnia liczba zabitych rocznie</b>	<b>średnia liczba rannych rocznie</b>
	<b>2001–2005</b>		
Ewaluowane	1,03	0,22	1,93
Referencyjne	0,81	0,33	1,29
<b>2007</b>			
Ewaluowane	0,83	0,00	1,33
Referencyjne	0,46	0,08	1,00

\* Za dodatni efekt netto uznajemy sukces, czyli zmniejszenie wypadkowości/liczby poszkodowanych, za ujemny efekt netto zaś porażkę, czyli zwiększenie wypadkowości/liczby ofiar lub ich zmniejszenie, ale słabsze na skrzyżowaniach ewaluowanych w porównaniu z referencyjnymi.

Tab. 1. Wyniki analiz dla 24 ewaluowanych skrzyżowań oraz dla skrzyżowań referencyjnych. Źródło: opracowanie własne

Na podstawie przedstawionych wartości wyliczono względny wzrost/spadek wypadkowości, zakładając, że będzie on interpretowalny empirycznie. Posługując się wzorem na przyrost względny ( $\frac{W_{2007} - W_{2001-2005}}{W_{2001-2005}} \times 100$ ), wyliczamy dane do tabeli 2.

<b>Wyniki</b>	<b>Wypadki</b>		
	<b>średnia wypadków rocznie</b>	<b>średnia liczba zabitych rocznie</b>	<b>średnia liczba rannych rocznie</b>
<b>Względny przyrost poziomu wypadkowości (w%)</b>			
Ewaluowane	-19,4	-100,0	-31,1
Referencyjne	-43,2	-75,8	-22,5

Tab. 2. Względny przyrost/spadek wypadkowości na badanych skrzyżowaniach Źródło: opracowanie własne

Po uwzględnieniu zaokrągleń otrzymujemy ostatecznie składowe efekty netto, wyliczone ze wzoru  $Wp_r - Wp_e$ , gdzie  $Wp_r$  oznacza względny przyrost poziomu wypadkowości dla skrzyżowań referencyjnych, zaś  $Wp_e$  – dla skrzyżowań ewaluowanych. Wyniki przekształcenia zawarto w tabeli 3.

<b>Wyniki</b>	<b>Składowe efekty netto (1) w %</b>		
	<b>średnia wypadków rocznie</b>	<b>średnia liczba zabitych rocznie</b>	<b>średnia liczba rannych rocznie</b>
Efekt netto	-23,8	24,2	8,6

Tab. 3. Wartości składowych efektu netto. Źródło: opracowanie własne.

Wyniki mają prostą empiryczną interpretowalność. Wartość  $-23,8\%$  oznacza, że w porównaniu ze skrzyżowaniami z grupy kontrolnej osiągnięto pogorszenie wyników, czyli *de facto* przebudowa przyczyniła się do zwiększenia wypadkowości. Jednocześnie warto jest zwrócić uwagę na dodatnie wartości składowych efektów netto w przypadku szacowania osób poszkodowanych, zwłaszcza jeśli chodzi o szacowaną liczbę osób zabitych. Można więc stwierdzić, że choć wypadkowość na ewaluowanych skrzyżowaniach spadła mniej niż na referencyjnych, liczba osób poszkodowanych w wypadkach spadła znacznie na skrzyżowaniach ewaluacyjnych niż referencyjnych. Co więcej, wynik ten nie jest spowodowany ogólnym wzrostem bezpieczeństwa samochodów czy koniecznością zapinania pasów, gdyż wówczas, jak można przypuszczać, podobne zjawisko obserwowalibyśmy w przypadku skrzyżowań referencyjnych i ostatecznie efekt netto byłby niższy (lub nie byłoby go w ogóle).

Jeżeli weźmiemy pod uwagę, że przedstawione liczby odzwierciedlają względny procent wzrostu/spadku danego zjawiska, który jest wynikiem interwencji, możemy procentowe wyniki efektu netto odnieść do statystyk zawartych w tabeli 1. Otrzymujemy wówczas efekt netto wyrażony w jednostkach (liczbie wypadków oraz liczbie osób rannych i zabitych), które zostały „zaoszczędzone” w wyniku interwencji. Otrzymany składowy efekt netto mnożymy razy wartość uzyskaną przez ewaluowane skrzyżowania w latach 2001–2005 [ $En_a\phi(2001 - 2005)_a$ ], gdzie  $En_a$  oznacza składowy efekt netto dla wskaźnika  $a$  (np. wypadkowości), zaś  $\phi(2001-2005)_a$  wartość danego wskaźnika (np. wypadkowości) w latach 2001–2005 dla ewaluowanych skrzyżowań. Wyniki przedstawiono w tabeli 4.

Wyniki	Składowe efekty netto (2)		
	średnia wypadków rocznie	średnia liczba zabitych rocznie	średnia liczba rannych rocznie
Efekt netto	$-24\% \times 1,03 = -0,245$ wypadku rocznie	$24\% \times 0,22 = 0,053$ osoby zabitej rocznie	$9\% \times 1,93 = 0,166$ osoby rannej rocznie

Tab. 4. Wartości składowych efektu netto w jednostkach. Źródło: opracowanie własne.

Tutaj też, jak w przypadku tabeli 2, ujemny efekt netto oznacza pogorszenie wyniku, dodatni zaś – polepszenie. Przekształcenie to ma tę zaletę, że pozwala interpretować nie względne wzrosty/spadki wypadkowości i odsetki osób poszkodowanych, lecz pozwala na interpretację w jednostkach mierzalnych pojedynczych wypadków i osób poszkodowanych. Z powyższego przekształcenia wynika, że w wyniku przebudowy zwiększono wypadkowość o około 1 wypadek na 4 lata, jak również zmniejszono liczbę ofiar śmiertelnych o jedną w ciągu 20 lat ( $0,053$  ofiary rocznie  $\times 20$  lat  $\approx 1$  ofiara śmiertelna).

Przekształcenie to pozwala również na zredukowanie liczby składowych efektów netto do dwóch tylko wskaźników: wzrostu/spadku wypadkowości *per se* i wzrostu/spadku liczby osób poszkodowanych. Wyniki zamieszczono w tabeli 5.

Wyniki	Składowe efekty netto (3a)	
	średnia wypadków rocznie	wskaźnik poszkodowanych w wypadkach (w przeliczeniu na osoby ranne)
Efekt netto (w jednostkach)	-0,245 wypadku rocznie	0,539 osób poszkodowanych rocznie

Tab. 5. Wartości składowych efektu netto (w przeliczeniu na osoby ranne). Źródło: opracowanie własne

Tak jak wspomniano powyżej, zdecydowano się nie łączyć ze sobą efektu netto związanego z wypadkowością i efektu netto związanego z liczbą osób poszkodowanych. Wziąwszy jednak pod uwagę dane zamieszczone w *Niebieskiej księdze* (2008), można założyć, że koszt zewnętrzny wypadku przypadający na 1 osobę zabita jest w przybliżeniu równy kosztowi zewnętrznemu 7 osób rannych i na tej podstawie wyliczyć syntetyczny wskaźnik kosztów zewnętrznych szkód na osobach. Zgodnie z tym przeliczeniem otrzymujemy dla średniej wypadkowości wynik tożsamy z poprzednio prezentowanym, a pozostałe dwa składowe efekty netto zostały przedstawione w sposób zagregowany. Wynik 0,539 oznacza, że dzięki przebudowie udało się uniknąć ponad dwóch osób rannych w wypadkach w skali dwóch lat.

Oczywiście te same dane można przedstawić także w przeliczeniu na osoby zabite, co pokazuje tabela 6.

Wyniki	Składowe efekty netto (3b)	
	średnia wypadków rocznie	wskaźnik poszkodowanych w wypadkach (w przeliczeniu na osoby zabite)
Efekt netto (w jednostkach)	-0,245 wypadku rocznie	0,077 osób poszkodowanych rocznie

Tab. 6. Wartości składowych efektu netto (w przeliczeniu na osoby zabite.) Źródło: opracowanie własne.

Wynik ten również otrzymano na podstawie opisanego powyższej przelicznika 1:7.

## 5. Interpretacja wyników

Interpretacja otrzymanych wyników nie nastręcza większych trudności. Ze względu na naukowy wymiar niniejszego projektu badawczego najistotniejszy jest fakt, że zastosowana metodyka pozwoliła na dokładny i wiarygodny dobór skrzyżowań podobnych, wielkość efektu netto istotnie różniła się zaś od efektu brutto, co uzasadnia konieczność szacowania tego pierwszego.

Natomiast z punktu widzenia ewaluacyjnego – oceny efektywności i skuteczności analizowanego projektu SPOT – bardzo ciekawy jest fakt, że

zaobserwowany efekt brutto również miał wartość ujemną, czyli że również na skrzyżowaniach referencyjnych zaobserwowano istotny spadek wypadkowości, zwłaszcza zaś spadek liczby zabitych. Wartości te są sprzeczne z ogólnymi trendami w panującymi w naszym kraju, w którym liczba ofiar wypadków drogowych jest od 2001 r. praktycznie na stałym poziomie (około 5500 zabitych), a liczba rannych (około 60 000 w 2001 r.) jedynie nieznacznie spada (o nie więcej niż 10%) (Symon 2008: 4).

Przyczyn takiego stanu rzeczy można upatrywać w kilku czynnikach.

Po pierwsze – część dobranych w niniejszym projekcie alternatywnych „czarnych punktów” mogła mieć charakter efemeryczny – były to miejsca, gdzie kilka wypadków zdarzyło się bez istotnej przyczyny (upodobniając je do czarnych punktów) i podobnie bez większej interwencji ustały. Tutaj jeszcze raz należy podkreślić, że problemem był fakt, że beneficjent w dokumentacji projektowej nie uzasadnił wyboru przebudowywanych skrzyżowań, spośród większej liczby rzeczywistych „czarnych punktów”, mogących łatwo stanowić skrzyżowania referencyjne. Problem ten można jednak również rozwiązać poprzez bardziej szczegółowe analizy na etapie późniejszych badań ewaluacyjnych (w skrajnym przypadku – poprzez oględziny wytypowanych wstępnie skrzyżowań referencyjnych).

Po drugie – skrzyżowania referencyjne również mogły zostać poddane innej interwencji, np. budowie obwodnicy danej miejscowości, budowie sygnalizacji świetlnej, budowie chodnika, oświetlenia czy postawieniu fotoradaru. Do tej uwagi odnosi się również postulat analiz eksperckich umieszczony w poprzednim punkcie, aczkolwiek należy się wówczas zastanawiać nad rzeczywistą zasadnością interwencji współfinansowanej z środków unijnych – być może istniejące problemy były możliwe do rozwiązania na szczeblu lokalnym czy krajowym (choć z drugiej strony unijna polityka transportowa zwraca uwagę na istotny problem wąskich gardeł).

Po trzecie – mogło się okazać, że kierowcy generalnie zaczęli uważać w punktach, w których niegdyś zdarzało się dużo wypadków, i „czarne punkty” podlegają również pewnym naturalnym przesunięciom.

Po czwarte – może się również okazać, że część istniejących czarnych punktów została wybrana nie tylko w sposób niesformalizowany, ale również błędnie, gdyż wśród przebudowywanych skrzyżowań znajdowały się takie, na których wypadkowość nie była duża (być może miały one ograniczoną przepustowość, jednak zwiększanie płynności ruchu nie było celem tego projektu). O tym może świadczyć fakt, że według bazy KGP w latach 2001–2005 na prawie połowie (11) z przebudowanych skrzyżowań zdarzyły się co najwyżej 2 wypadki<sup>1</sup>.

Po piąte wreszcie – analizowany okres po przebudowie był krótki i chociaż dotyczył łącznie 24 lub 23 obiektów (a zatem odpowiadał prawdopodobieństwu wystąpienia wypadku na jednym skrzyżowaniu w ciągu ćwierci wieku), to jednak wciąż możliwe jest zniekształcenie wyników, ze względu na rzadkość wybranych zdarzeń.

Należy zauważyć, że zaproponowany model szacuje prawdopodobieństwo wybrania skrzyżowania do projektu tylko na podstawie dość ograniczonego zestawu zmiennych niezależnych opisanych we wcześniejszej części tekstu. Nie bierze jednak pod uwagę zmiennych, które mogły istotnie wpłynąć na ostateczną wypadkowość na ewaluowanych skrzyżowaniach, a których wpływ jest często niemierzalny bądź w ogóle ukryty. Mamy tu na myśli przede wszystkim kwestię przyzwyczajenia kierowców oraz rozmaite zdarzenia o charakterze intencjonalnym, które w żadnej bazie danych nie figurują.

Oczywiste jest, że zarówno zmienne słabo mierzalne, jak i ukryte mogą mieć większy wpływ na skrzyżowania przebudowane niż na skrzyżowania referencyjne właśnie z tego powodu, że te pierwsze uległy przeobrażeniom, a drugie nie. Są to ograniczenia modelu, z których trzeba zdawać sobie sprawę, które jednak nie sprawiają, że otrzymane wartości efektu netto są empirycznie nieoszacowywalne. Otrzymane i opisane wartości składowych efektów netto zostały opracowane na podstawie pewnych dobrze określonych i łatwych do zmierzenia zmiennych i na tej podstawie można mówić o sukcesach lub porażkach ocenianych projektów.

## 6. Wnioski

Zastosowanie metodologii ewaluacji wpływu interwencji współfinansowanych z środków unijnych jest narzędziem pozwalającym na rzetelną ocenę rzeczywistych efektów realizowanych projektów na podstawie wybranych kryteriów. Niemniej adaptacja tej metodologii do oceny efektów netto inwestycji w infrastrukturę drogową wskazuje na kilka istotnych kwestii, które powinny być wzięte pod uwagę przy podejmowaniu podobnych badań.

Przed wszystkim metoda pomija inne niż założone przez ewaluatorów efekty danej interwencji (np. tzw. efekty uboczne czy wartości dodane, nieprzewidziane przez autorów projektów). Rozważyć należałoby zakres użyteczności wyników badań efektu netto i ich właściwą interpretację – przecież ujemny wynik efektu netto przebudowanych skrzyżowań nie powinien stać się argumentem przemawiającym za zatrzymaniem takich inwestycji w przyszłości. Wyniki badań pozwalają określić również słabe strony metodologii. Poważnym ograniczeniem okazało się wielowymiarowe otoczenie prowadzonych inwestycji i wynikający z tego wpływ czynników niemierzalnych lub ukrytych – trudnych, a nawet niemożliwych do wzięcia pod uwagę przy opracowywaniu metodologii ewaluacji wpływu.

Niemniej należy podkreślić, że zastosowanie omawianej w niniejszym artykule metodologii powinno przyczynić się do podniesienia jakości procesu zarządzania, wdrażania, monitorowania, oceny i kontroli Narodowej Strategii Spójności realizowanej w bieżącym okresie budżetowania. Prawidłowa ocena wpływu inwestycji infrastrukturalnych jest tym bardziej istotna, że w obecnej chwili w Komisji Europejskiej toczą się rozmowy na temat przyszłości polityki spójności po 2013 r. Przed rozpoczęciem projektowania

budżetu – jego wielkości oraz kierunków strategicznych wsparcia, Komisja Europejska potrzebuje rzetelnych informacji na temat efektów dotychczasowych interwencji. Rozwój metodologii ewaluacji wpływu, a następnie ocena efektów netto mogą się więc okazać cenną pomocą w generowaniu argumentów do dyskusji na temat kształtu polityki spójności w przyszłości.

### Informacje o autorach

**Dr Dominika Wojtowicz** – Katedra Europeistyki Akademia Leona Koźmińskiego.  
E-mail: doma@kozminski.edu.pl.

**Dr Michał Wolański** – Szkoła Główna Handlowa w Warszawie.  
E-mail: michal.wolanski@sgh.waw.pl

### Przypisy

<sup>1</sup> Należy tutaj podkreślić niedoskonałość istniejącej bazy – dla potrzeb studiów wykonano opracowanie bardziej szczegółowe dane, wraz z liczbą kolizji, jednakże brak było porównywalnych danych dla skrzyżowań referencyjnych. Niedoskonałość ta może być ograniczona wraz ze wprowadzeniem Systemu Ewidencji Wypadków i Kolizji SEWIK.

### Bibliografia

- Agrosynergie Groupement Européen d'Intérêt Economique 2008. *Ocena uruchomienia płatności bezpośrednich z tytułu upraw owoców i warzyw według modelu regionalnego*.
- Babbie, E. 2007. *Badania społeczne w praktyce*, Warszawa: Wyd. Naukowe PWN.
- Boarnet, M. 2006. *Conducting Impact Evaluations in Urban Transport. Doing Impact Evaluation Series*, nr 5, Thematic Group on Poverty Analysis, Monitoring and Impact Evaluation, The World Bank.
- Center for Global Development 2006. *When Will We Ever Learn? Improving Lives Through Impact Evaluation*, Evaluation Gap Working Group, Centre for Global Development, New York.
- Górniak, J. 2007. *Ewaluacja w cyklu polityk publicznych*, w: Mazur, S. (red.) *Ewaluacja funduszy strukturalnych – perspektywa regionalna*, Kraków: Uniwersytet Ekonomiczny w Krakowie Małopolska Szkoła Administracji Publicznej.
- Komisja Europejska 1997. *Ewaluacja programów wydatków Unii Europejskiej. Przewodnik Ewaluacja w połowie okresu i ex-post XIX/02 – Synthèse budgétaire et évaluation*, Dyrekcja Generalna XIX – Budżet.
- Konarski, R. i M. Kontarowski 2007. *Zastosowanie metody propensity score matching w ewaluacji ex-post*, w: Haber, A. (red.) *Ewaluacja ex-post. Teoria i praktyka badawcza*, Warszawa: PARP.
- Michalek, J. 2007. *Construction and the measurement of the Rural Development Index. Application to evaluation of RD policies. Part II*, Conceptual Reports CR3-2, University of Kiel.
- Michalek, J i N. Zarnekow 2008. *Empirical measurement of quality of life in rural areas in Slovakia using the RDI Index*, Regional Report RR3-1, University of Kiel.
- Michalek, J., Zarnekow, N., Stulrajter, Z. i L. Klank 2008. *Cartographic and statistical analysis of RDI in Poland and Slovakia*, Regional Report RR3-2, University of Kiel.
- Niebieska księga, Infrastruktura drogowa*. 2008. Warszawa: Jaspers.
- Ocena uruchomienia płatności bezpośrednich z tytułu upraw owoców i warzyw według modelu regionalnego*. 2008. Agrosynergie Groupement Européen d'Intérêt Economique.

- Olejniczak, K., Kozak, M. i B. Ledzion 2008. *Teoria i praktyka ewaluacji interwencji publicznych. Podręcznik akademicki*, Warszawa: Wyd. Akademickie i Profesjonalne Akademii Leona Koźmińskiego.
- Rettore, E., Trivellato, U. i A. Martini 2003. La valutazione delle politiche del lavoro in presenza di selezione: migliorare la teoria, i metodi o i dati? *Politica economica. Rivista di studi e ricerche per la politica economica*, nr 3, s. 301–342.
- Strawiński, P. 2008. Quasi-eksperymentalne metody ewaluacji, w: *Środowisko i warsztat ewaluacji*, Warszawa: PARP.
- Symon, E. 2008. *Wypadki drogowe w Polsce w 2007 roku*, Warszawa: KGP.
- Walle, D. 2009. Impact Evaluations of Rural Road Projects. *Journal of Development Effectiveness*, vol. 1, nr 1.
- White, H. 2006. *Impact Evaluation: The Experience of the Independent Evaluation Group of the World Bank*, Washington: World Bank.