



KOMISJA
EUROPEJSKA

Bruksela, dnia XXX r.
SWD(2014) XXX

DOKUMENT ROBOCZY SŁUŻB KOMISJI

Wspólne metody oceny pomocy państwa

DOKUMENT ROBOCZY SŁUŻB KOMISJI

Wspólne metody oceny pomocy państwa

Spis treści

1	Wprowadzenie	2
2	Cele oceny pomocy państwa	3
3	Plan oceny	5
3.1	Cele programu pomocy, które należy poddać ocenie	5
3.2	Pytania w ramach oceny	5
3.3	Wskaźniki rezultatu	6
3.4	Metody: znalezienie odpowiedniej podstawy porównania	7
3.5	Gromadzenie danych: korzystanie z najlepszych możliwych źródeł	9
3.6	Harmonogram oceny	11
3.7	Organ przeprowadzający ocenę: zapewnianie niezależności i wiedzy fachowej	12
3.8	Jawność: ułatwianie zaangażowania zainteresowanych stron	13
4	Kryteria wyboru dotyczące programów pomocy, które mają zostać poddane ocenie	14
4.1	Duże programy pomocy, w tym programy objęte przepisami ogólnego rozporządzenia w sprawie wyłączeń grupowych	14
4.2	Nowatorskie programy pomocy	15
4.3	Programy pomocy, na które oddziałują znaczące przewidziane zmiany	15
4.4	Inne programy pomocy	16
Załącznik I: Dodatek techniczny dotyczący odpowiednich metod określania wpływu przyczynowego		17
Załącznik II: Wykaz możliwych wskaźników rezultatu		39
Załącznik III: Glosariusz		44
Załącznik IV: Dokumenty referencyjne		45

1 Wprowadzenie

Państwa członkowskie udzielają pomocy państwa, aby ułatwić osiągnięcie różnego rodzaju celów polityki, np. zmniejszyć różnice regionalne w kraju, wspierać działalność badawczą, rozwojową i innowacyjną lub promować wysoki poziom ochrony środowiska.

Unijne zasady pomocy państwa, w których określono rodzaje pomocy zgodne ze wspólnym rynkiem, opierają się na systemie działań kontrolnych prowadzonych *ex ante*: programy pomocy¹ zatwierdza się na podstawie wcześniej określonych kryteriów oceny przy założeniu, że jeżeli programy są zgodne z kryteriami oceny, to ich pozytywne skutki przeważają nad wszelkimi skutkami negatywnymi. Ocena programów odbywa się zazwyczaj bez odpowiedniej oceny ich rzeczywistego wpływu na rynki na przestrzeni czasu.

Do tej pory, stosując unijne zasady pomocy państwa, przywiązywano stosunkowo niewielkie znaczenie do dowodów *ex post* świadczących o rzeczywistych rezultatach osiągniętych dzięki środkom publicznym czy o wpływie pomocy państwa na konkurencję. Jednak decydenci – zarówno na szczeblu państwa członkowskiego, jak i na poziomie UE – powinni uwzględniać wymierne wyniki pomocy państwa przyznanej w przeszłości oraz wyciągnięte z nich wnioski. Dzięki temu programy finansowane w ramach pomocy państwa będą skuteczniejsze i będą powodować mniej zakłóceń na rynkach; zwiększy się także wydajność przyszłych programów i ewentualnie przyszłych zasad przyznawania pomocy państwa.

Wiele państw dokonało już oceny swoich środków pomocowych, nawet jeżeli nie zawsze przeprowadzano oceny regularnie². Również wydatki UE (w tym finansowanie z unijnych funduszy strukturalnych i inwestycyjnych takich jak EFRR, EFS i EFRROW) podlegają ocenie *ex ante*, ocenie bieżącej i ocenie *ex post* zgodnie z obowiązującymi przepisami i wytycznymi opublikowanymi przez Komisję³. Aby uniknąć powielania ocen prowadzonych przez państwa członkowskie, w wytycznych „Koncepcje i zalecenia” dotyczących monitorowania i oceny wyjaśniono, że wymogi w zakresie oceny europejskich funduszy strukturalnych i inwestycyjnych można spełnić, przeprowadzając oceny wymagane w myśl zasad pomocy państwa.

1 W ramach programów udziela się większości przyznawanej pomocy: według danych z tablicy wyników z 2013 r. zatwierdzone programy pomocy stanowią 23 % wszystkich środków pomocy i odpowiadają za 55 % kwot pomocy; szereg programów objętych wyłączeniami grupowymi stanowi 63 % wszystkich środków pomocy i odpowiada za około 32 % kwot pomocy. Zgodnie z rozporządzeniem Rady nr 659/1999 „program pomocy” oznacza „każde działanie, na którego podstawie, bez dalszego wprowadzania w życie wymaganych środków, można dokonać wypłat pomocy indywidualnej na rzecz przedsiębiorstw określonych w ustawie w sposób ogólny i abstrakcyjny oraz każde działanie, na którego podstawie pomoc, która nie jest związana z konkretnym projektem, może zostać przyznana jednemu lub kilku przedsiębiorstwom na czas nieokreślony i/lub w nieokreślonej kwocie”.

2 Na przykład w kilku państwach członkowskich regularnie sporządza się sprawozdania z oceny pomocy państwa dla Trybunału Obrachunkowego lub Parlamentu.

3 W wytycznych Komisji dotyczących oceny za okres finansowania 2014–2020 (dostępnych pod adresem: http://ec.europa.eu/regional_policy/information/evaluations/guidance_en.cfm) określono szczegółowo odpowiednie koncepcje i zalecenia.

Inicjatywa unowocześnienia polityki w dziedzinie pomocy państwa⁴ ma na celu skoncentrowanie działań Komisji w ramach egzekwowania prawa na większych programach pomocy, które mogą mieć największy wpływ na wspólny rynek. Jednocześnie analiza przypadków o charakterze bardziej lokalnym, mających niewielki lub bardziej ograniczony wpływ na wymianę handlową, zostanie uproszczona, w tym dzięki zapewnieniu państwom członkowskim większej elastyczności pod względem wdrażania takich środków pomocy poprzez poszerzenie zakresu nowego ogólnego rozporządzenia w sprawie wyłączeń grupowych⁵. W celu zapewnienia, aby – ogólnie rzecz biorąc – pozytywne skutki pomocy państwa (pod względem osiągnięcia swojego pierwotnego celu) w dalszym ciągu przeważały nad potencjalnym negatywnym wpływem na konkurencję i wymianę handlową, oraz w celu uniknięcia nadmiernego zakłócenia rynku większe uproszczenie należy połączyć z większą przejrzystością, wzmocnioną kontrolą zgodności z zasadami pomocy państwa na szczeblu krajowym i unijnym, a także ze skuteczną oceną⁶.

W niniejszym dokumencie określono wspólne metody oceny programów pomocy państwa. Ma on w celu udzielenie wskazówek organom publicznym biorącym udział w procesie planowania i przeprowadzania ocen.

2 Cele oceny pomocy państwa

Ogólnym celem oceny pomocy państwa jest ocena względnych pozytywnych i negatywnych skutków programu, tj. publicznego celu pomocy pod względem wpływu na konkurencję i wymianę handlową między państwami członkowskimi. Za pomocą oceny pomocy państwa można wyjaśnić, czy i w jakim stopniu pierwotne cele programu pomocy zostały osiągnięte (tj. ocenić pozytywne skutki) oraz określić wpływ programu na rynki i konkurencję (tj. ewentualne negatywne skutki). Zatem w swoim zamierzeniu ocena różni się od dwóch działań *ex post* prowadzonych obecnie przez Komisję w odniesieniu do programów pomocy państwa – tj. monitorowania⁷ i sprawozdawczości⁸.

Ocena pomocy państwa powinna w szczególności umożliwić ocenę bezpośredniego efektu zachęty pomocy wobec beneficjenta (tj. czy w wyniku przyznanej pomocy beneficjent obrał

4 Komunikat Komisji do Parlamentu Europejskiego, Rady, Europejskiego Komitetu Ekonomiczno-Społecznego, Komitetu Regionów: Unowocześnienie unijnej polityki w dziedzinie pomocy państwa, 8.5.2012, COM(2012) 209 final.

5 Rozporządzenie Komisji (UE) nr .../2014 z dnia XXX r. uznające niektóre rodzaje pomocy za zgodne z rynkiem wewnętrznym w zastosowaniu art. 107 i 108 Traktatu.

6 Zob. także konkluzje Rady w sprawie reformy kontroli pomocy państwa z dnia 13 listopada 2012 r.

7 Monitorowanie prowadzone przez Komisję oznacza okresowy przegląd legalności próbek środków pomocy państwa wdrażanych przez państwa członkowskie. Ma to zapewnić prawidłowe wykonanie decyzji Komisji przez państwa członkowskie, a także przestrzeganie przez państwa członkowskie odpowiednich przepisów prawnych (tj. przepisów zawartych w ogólnym rozporządzeniu w sprawie wyłączeń grupowych). Komisja ocenia także przestrzeganie zasad i warunków *ex ante* wśród reprezentatywnej grupy spraw.

8 Głównym celem sprawozdań rocznych składanych przez państwa członkowskie jest zwiększenie przejrzystości pomocy państwa przyznanej przez państwa członkowskie. Stanowią one również źródło wiarygodnych statystyk na potrzeby kształtowania polityki i monitorowania. Dane zawarte w sprawozdaniach rocznych dostarczają głównie informacji ilościowych (np. przedstawiających cele, na jakie przeznaczono pomoc państwa, i przy jakim poziomie budżetu). Komisja korzysta ze sprawozdań państw członkowskich w celu przygotowania tablicy wyników w dziedzinie pomocy państwa.

inny kierunek działań oraz jak istotny wpływ miała przyznana pomoc). Dokonując oceny, należy również szacunkowo określić ogólne pozytywne i negatywne skutki programu pomocy w odniesieniu do osiągnięcia założonego celu polityki oraz w odniesieniu do konkurencji i wymiany handlowej; można także zbadać proporcjonalność i odpowiedniość wybranego instrumentu pomocy.

W oparciu o ocenę można potwierdzić, czy założenia stanowiące podstawę wcześniejszego zatwierdzenia programu pomocy są nadal aktualne i mogą przyczynić się do lepszego opracowania przyszłych programów pomocy i zasad regulujących pomoc państwa. Ocena może też posłużyć jako podstawa do skorygowania przyszłych interwencji państwa w celu zwiększenia skuteczności i wydajności pomocy w zakresie niezbędnym do zagwarantowania, że pozytywne skutki są wystarczające, aby uzasadnić akceptację zakłócenia rynku spowodowanego przez interwencję. Tego rodzaju ulepszenia przyszłych programów mogą obejmować zarówno korekty ich struktury, w tym zmiany kryteriów wyboru i bardziej wnikliwą ocenę efektu zachęty, jak i bardziej istotne zmiany, takie jak promowanie stosowania alternatywnej formy pomocy, ponowne zdefiniowanie celów lub beneficjentów docelowych lub rozważenie opcji niezwiązanych z pomocą w celu osiągnięcia tych samych celów polityki.

Należy ustalić właściwy harmonogram oceny, gwarantujący odpowiednią ilość czasu na zebranie wystarczających dowodów i umożliwiający jak najszybsze przedstawienie wyników decydentom, aby potencjalne ulepszenia można było wprowadzić w odpowiednim czasie⁹. W związku z tym oceny pomocy państwa należy zwykle traktować jako oceny bieżące, prowadzone w trakcie realizacji programu pomocy, a nie tylko jako oceny *ex post* prowadzone dopiero po zakończeniu realizacji programu. Należy uwzględnić przypadki szczególne, w których pełne skutki interwencji będą widoczne po dłuższym czasie, a ocena umożliwia tylko oddanie i zmierzenie skutków początkowych.

W ostatecznym rozrachunku ocena pomocy państwa powinna stanowić źródło wiedzy zarówno dla Komisji, jak i dla państw członkowskich. Aby było to możliwe, ocena powinna spełniać pewne minimalne normy jakości. Komisja powinna zatem zapewnić przeprowadzenie odpowiedniej kontroli jakości ocen. W szczególności Komisja dokona szczegółowej analizy ogólnej wiarygodności oceny i wykaże ewentualne braki na dwóch zasadniczych etapach, a mianowicie w planie oceny i sprawozdaniu końcowym. W stosownych przypadkach Komisja może zwrócić się o pomoc do niezależnych ekspertów zewnętrznych, aby uczestniczyli oni w kontroli jakości oceny.

Komisja może również zorganizować sesje szkoleniowe i warsztaty dla organów administracji krajowych w zakresie metod i technik stosowanych podczas oceny. Ponadto istnieje możliwość wymiany pozytywnych doświadczeń i najlepszych praktyk państw członkowskich w celu opracowania skuteczniejszych programów pomocy w przyszłości.

9 W niektórych wytycznych dotyczących pomocy państwa mowa o normalnym czasie trwania obejmującym cztery lata dla ocenianych programów pomocy.

Korzyści wynikające z przeprowadzenia ocen będą widoczne za kilka lat, gdy pierwsze sprawozdania z oceny będą gotowe, a zawarte w nich wnioski i zalecenia zostaną udostępnione. Wówczas będzie można z nich skorzystać w celu ulepszenia struktury kolejnych programów pomocy i ewentualnie zasad regulujących pomoc państwa. W perspektywie średnio- i długoterminowej ocena może stopniowo prowadzić do bardziej fundamentalnych zmian w ogólnym podejściu do pomocy państwa.

3 Plan oceny

Kompleksowy plan oceny programu pomocy państwa należy sporządzić na wczesnym etapie, równoległe z opracowaniem programu. Zatwierdzenie przez Komisję planu oceny jest konieczne, by zapewnić równe traktowanie. Plan należy następnie rygorystycznie wdrożyć.

Na ogół uznaje się, że oceny są skuteczniejsze, jeżeli zostały właściwie zaplanowane i przygotowane z wyprzedzeniem, w szczególności ponieważ ułatwia to gromadzenie odpowiednich danych. Istnieje także prawdopodobieństwo, że dzięki wczesnemu planowaniu znacznemu zmniejszeniu ulegną zasoby wymagane do przeprowadzenia oceny i że ostatecznie poprawi się jej jakość.

Plan oceny zgłoszony Komisji przez państwo członkowskie zgodnie z odpowiednimi przepisami powinien zawierać co najmniej następujące minimalne elementy.

3.1 Cele programu pomocy, które należy poddać ocenie

Pierwszy etap oceny programu polega na wyraźnym ustaleniu „logiki interwencji” leżącej u podstaw programu pomocy, z opisaniem potrzeb i problemów, których program ma dotyczyć, beneficjentów docelowych i inwestycji, celów ogólnych i szczegółowych programu oraz oczekiwanego wpływu. Należy także wspomnieć główne założenia dotyczące czynników zewnętrznych, które mogą mieć wpływ na program.

3.2 Pytania w ramach oceny

W planie oceny należy określić zakres oceny, tzn. plan powinien zawierać konkretne pytania, na które można odpowiedzieć, przedstawiając dane ilościowe i niezbędne potwierdzające dowody. Pytania w ramach oceny powinny dotyczyć głównie wpływu programu pomocy państwa. Pytania te można sklasyfikować według trzech następujących poziomów:

1. Bezpośredni wpływ pomocy na beneficjentów, na przykład:

- Czy pomoc miała istotny wpływ na przebieg działań podjętych przez beneficjentów pomocy (efekt zachęty)?
- Czy pomoc miała wpływ na sytuację beneficjentów (na przykład czy uległa zmianie ich pozycja konkurencyjna lub ryzyko niewykonania zobowiązania)?

- W jakim stopniu pomoc przyniosła oczekiwane skutki?
- Czy beneficjenci w różny sposób odczuli wpływ pomocy (na przykład w zależności od wielkości, lokalizacji lub sektora)?

2. Pośredni wpływ programu pomocy, na przykład:

- Czy program miał efekt rozlania w odniesieniu do działalności innych przedsiębiorstw lub innych regionów geograficznych? Czy pomoc spowodowała wyparcie inwestycji dokonywanych przez innych konkurentów lub przyciągnięcie działalności z sąsiednich regionów?
- Czy program przyczynił się do osiągnięcia odpowiedniego celu polityki?
- Czy możliwe jest zmierzenie zagregowanych skutków w odniesieniu do konkurencji i wymiany handlowej?

3. Proporcjonalność i odpowiedniość programu pomocy, na przykład:

- Czy pomoc była proporcjonalna do problemu, którego dotyczyła? Czy możliwe byłoby osiągnięcie takich samych skutków przy mniejszej pomocy lub innej formie pomocy (na przykład pożyczki zamiast dotacji)?
- Czy wybrano najskuteczniejszy instrument pomocy? Czy inne instrumenty pomocy lub rodzaje interwencji byłyby bardziej odpowiednie do osiągnięcia danego celu?

Ocena powinna w miarę możliwości obejmować wpływ programu pomocy na wszystkich trzech poziomach, uwzględniając odpowiednie pytania dotyczące celów programu. Zazwyczaj jednak bezpośredni wpływ pomocy na beneficjentów jest tym rodzajem wpływu, który można zbadać najdokładniej. W praktyce większość opracowanych metod oceny służy do zmierzenia tego rodzaju wpływu. Ponadto ocena bezpośrednich skutków pomocy, w tym efektu zachęty, ma największą wagę, ponieważ może stanowić źródło cennych informacji na temat rodzajów skutków pośrednich i oczekiwanych zakłóceń. W szczególności jeżeli pomoc nie zapewnia efektu zachęty, można założyć, że pomoc zakłóca konkurencję w takim sensie, że zapewnia danym beneficjentom nieoczekiwane zyski.

3.3 Wskaźniki rezultatu

Pytania w ramach oceny powinny prowadzić do wyboru określonych wskaźników rezultatu, które oddają ilościowe informacje dotyczące wyników osiągniętych w ramach programu pomocy państwa. Załącznik II zawiera orientacyjny i niewyczerpujący wykaz wskaźników rezultatu, obejmujący zarówno bezpośredni, jak i pośredni wpływ programu, w tym ewentualny wpływ na konkurencję i wymianę handlową. Wskaźniki rezultatu będą zależały od celu ocenianej pomocy. W planie oceny należy wyjaśnić, dlaczego wybrane wskaźniki odgrywają najważniejszą rolę w mierzeniu wpływu danego programu pomocy.

3.4 Metody: znalezienie odpowiedniej podstawy porównania

W ramach ocen pomocy państwa powinna istnieć możliwość określenia przyczynowego wpływu samego programu, niezakłócanego przez inne zmienne, które mogły mieć wpływ na zaobserwowane wyniki, np. ogólne warunki makroekonomiczne lub niejednorodność przedsiębiorstwa (np. różnice w wielkości przedsiębiorstwa, jego lokalizacji, środkach finansowych lub zdolnościach zarządzania). W planie oceny należy określić główne metody, które zostaną wykorzystane do określenia skutku pomocy, i omówić, dlaczego dane metody prawdopodobnie są odpowiednie dla danego programu.

Przyczynowy wpływ jest to różnica między wynikami osiągniętymi w przypadku przyznania pomocy a wynikami osiągniętymi w przypadku jej braku. Chociaż wyniki uzyskane w przypadku przyznania pomocy bada się w odniesieniu do przedsiębiorstw, które otrzymały pomoc, wyniki w przypadku braku pomocy są mierzone jedynie w odniesieniu do przedsiębiorstw, które nie otrzymały pomocy. Z definicji nie bada się, jakie byłyby wyniki, gdyby przedsiębiorstwom, które otrzymały pomoc, jej nie przyznano. Aby oszacować wpływ pomocy na jej beneficjenta, konieczne jest zatem stworzenie scenariusza alternatywnego w oparciu o najbardziej porównywalne przedsiębiorstwo(-a) lub grupę kontrolną.

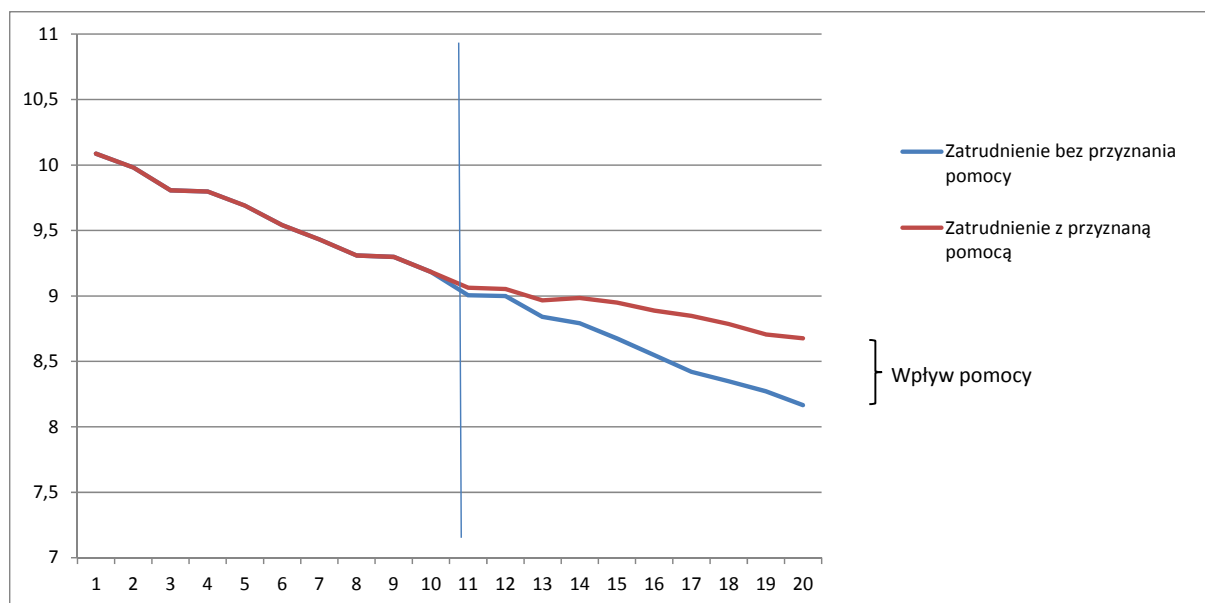
Jakość danej grupy kontrolnej ma kluczowe znaczenie dla wiarygodności oceny.

Sytuacja przedsiębiorstw, które otrzymały pomoc, może różnić się od sytuacji przedsiębiorstw, które nie otrzymały pomocy. Przedsiębiorstwa mogą na przykład zmagać się z różnymi lokalnymi warunkami dotyczącymi podaży i popytu, mieć trudniejszy dostęp do kredytu lub być bardziej lub mniej skuteczne. Te czynniki mogą mieć wpływ na wyniki działalności lub poziom działalności przedsiębiorstw, zarówno w przypadku otrzymania pomocy, jak i w przypadku jej braku. Porównanie wyników działalności beneficjentów z wynikami podmiotów niebędących beneficjentami może odzwierciedlić ten stan rzeczy w większym stopniu niż sam wpływ pomocy. Ocena programu pomocy nie może opierać się zatem na prostym porównaniu beneficjentów i podmiotów niebędących beneficjentami, ale musi uwzględniać różne cechy tych dwóch grup przedsiębiorstw, zarówno te, które można, jak i te, których nie można obserwować.

Na przykład w przypadku pomocy regionalnej beneficjenci pomocy w regionach, w których panują niekorzystne warunki rynkowe (tj. w których produkt lokalny, rynek pracy lub rynek kapitałowy są słabe), zazwyczaj osiągają gorsze wyniki niż podmioty niebędące beneficjentami w lepiej prosperujących regionach. W żaden sposób nie odzwierciedla to jednak wpływu samej pomocy. Istotną kwestią jest pytanie, czy osiągnęli lepsze wyniki niż te, które uzyskaliby w przypadku braku pomocy, a nie pytanie, czy osiągnęli lepsze wyniki niż podmioty niebędące beneficjentami w innych regionach.

Podczas określania wpływu pomocy należy również uwzględnić ogólne tendencje w branży. Nawet jeżeli beneficjenci pomocy regionalnej ograniczą liczbę swoich pracowników, pomoc w dalszym ciągu może być skuteczna. Na przykład gdy warunki w danej branży jako całości ulegają pogorszeniu, a wszystkie przedsiębiorstwa redukują liczbę miejsc pracy, beneficjenci

pomocy mogą zredukować zatrudnienie w mniejszym stopniu niż w przypadku, gdyby nie otrzymali pomocy. Taką sytuację ilustruje poniższy wykres, na którym przedstawiono tendencję spadkową liczby miejsc pracy w przedsiębiorstwach, które otrzymały pomoc, zarówno przed, jak i po przyznaniu pomocy. Jednak po otrzymaniu pomocy przez dane przedsiębiorstwo tendencja spadkowa ulega zmniejszeniu. Różnica między przedłużoną linią ukazującą tendencję w przypadku nieprzyznania pomocy a linią ukazującą rzeczywiste zatrudnienie oferowane przez dane przedsiębiorstwo po otrzymaniu pomocy stanowi efekt pozytywnego wpływu pomocy.



Rysunek 1 – pozytywny wpływ pomocy w przypadku negatywnej tendencji bieżącej

Gdy podmioty niebędące beneficjentami samodzielnie decydowały, czy ubiegać się, czy nie ubiegać się o pomoc, pojawia się szczególny problem dotyczący identyfikacji grupy kontrolnej. Na przykład jeżeli kwalifikują się wszystkie przedsiębiorstwa (tj. wszystkie przedsiębiorstwa, które przedstawiają projekt i ubiegają się o pomoc, otrzymują pomoc), wówczas przedsiębiorstwa, które nie ubiegają się o pomoc, mogą być przedsiębiorstwami bez projektów. Wyniki przedsiębiorstw mogą wskazywać, że przedsiębiorstwa, które nie otrzymały pomocy, radziły sobie gorzej w ujęciu bezwzględnym i względnym niż przedsiębiorstwa, które otrzymały pomoc. Wniosek ten można jednak całkowicie wyjaśnić faktem, że ta druga grupa – w przeciwieństwie do pierwszej – posiadała projekt; oznacza to, że kierownictwo pierwszej grupy nie wykazuje zainteresowania lub brak mu kreatywności. Istotne jest zatem, aby przedsiębiorstwa w grupie kontrolnej (przedsiębiorstwa, które nie skorzystały z pomocy) były częścią tej grupy z powodów, które nie mają wpływu na zmierzone wyniki. W szczególności jeżeli przedsiębiorstwa samodzielnie i dobrowolnie postanowiły, że nie będą ubiegać się o pomoc, warunek ten może nie zostać spełniony.

Wszelkie systematyczne różnice między beneficjentami pomocy państwa a podmiotami niebędącymi beneficjentami należy odpowiednio uwzględnić w projekcie oceny, aby uniknąć błędu systematycznego w wynikach (obciążenie selekcji). W ostatnich dekadach opracowano

kilka rzetelnych metod, aby rozwiązać ten problem. Wybór metody zależy od struktury danego programu pomocy państwa i dostępnych danych. Wszystkie metody mają swoje ograniczenia i sprawdzają się tylko w przypadku, gdy aktualne są pewne założenia. Uznanie i otwarte omówienie tych ograniczeń i założeń ma kluczowe znaczenie dla wiarygodności badania.

Zapewnienie losowości procesu wykorzystywanego do selekcji beneficjentów jest jednym ze sposobów zagwarantowania, że ocena nie jest tendencyjna. Jeżeli beneficjentów pomocy wybiera się całkowicie w sposób losowy, wszelkie różnice systematyczne zaobserwowane w wynikach działalności przedsiębiorstw można przypisać pomocy. Metoda ta może być jednak trudna do zastosowania w praktyce, w szczególności w przypadku dużych istniejących programów. Inne metody mają na celu wykorzystanie istniejących źródeł zróżnicowania zewnętrznego w otoczeniu, w którym funkcjonują przedsiębiorstwa (tj. zróżnicowania niezależnego od parametrów i zmiennych w modelu), w celu określenia związku przyczynowo-skutkowego¹⁰. W załączniku I do niniejszych wytycznych przedstawiono bardziej szczegółowo najodpowiedniejsze metody, koncentrując się na praktycznych aspektach ich stosowania. Omówiono w nim sposób, w jaki każda metoda określa związek przyczynowo-skutkowy, co jest szczególnie istotne w kontekście ocen pomocy państwa, w których projekt *ex ante* oceny służy do zapewnienia możliwości dokonywania właściwej oceny skutków pomocy.

Ponadto należy skontrolować wpływ wielu rodzajów pomocy niezależnie od tego, czy pomoc pochodzi z jednego programu, wielu programów czy też przyznano ją jako pomoc *ad hoc*. Jeżeli podmioty niebędące beneficjentami w danym programie otrzymają pomoc z innych programów lub jeżeli beneficjenci danego programu otrzymają dodatkową pomoc z innych programów, ocena skutków danego programu pomocy prawdopodobnie zostanie zniekształcona.

3.5 Gromadzenie danych: korzystanie z najlepszych możliwych źródeł

Należy gromadzić spójne i wystarczające dane zarówno na temat beneficjentów pomocy, jak i grupy kontrolnej. Identyfikowanie wymaganych danych i uzyskiwanie dostępu do źródeł formularzy danych stanowią część planowania oceny.

Skuteczne monitorowanie interwencji oraz dokładne gromadzenie i przetwarzanie danych mają zasadnicze znaczenie dla zapewniania jakości oceny. W związku z tym po zatwierdzeniu programu pomocy należy niezwłocznie wprowadzić mechanizm monitorowania interwencji oraz gromadzenia i przetwarzania odpowiednich danych. Mechanizm taki prawdopodobnie pozwoli znacznie obniżyć koszty oceny.

Upewnienie się, czy zgromadzone zostały niezbędne dane dotyczące wnioskodawców ubiegających się o pomoc i beneficjentów pomocy, stanowi kluczowy element

10 Do najczęściej stosowanych metod należą różnica w różnicach, projekt nieciągłości regresji i zmienne instrumentalne.

opracowywania planu oceny, w przypadku gdy dostępność tych danych może stanowić jeden z warunków kwalifikujących do pomocy.

Z wyjątkiem danych dotyczących wnioskodawców ubiegających się o pomoc (w tym w stosownych przypadkach wnioskodawców, których wnioski zostały odrzucone) źródła danych dotyczących beneficjentów pomocy i dotyczących grupy kontrolnej muszą być takie same, aby dane te były porównywalne. Wysoce prawdopodobne jest, że dane będą musiały zostać pozyskane z różnych źródeł, np. dane z baz danych zawierających informacje na temat odbiorców pomocy zostaną połączone z danymi pochodzącymi z rejestrów handlowych. Ocena może wymagać korzystania z istniejących źródeł danych, takich jak administracyjne źródła danych (np. urząd skarbowy, rejestr handlowy, badania na temat innowacji i urząd patentowy). Dlatego też w planie oceny należy przeprowadzić przegląd istniejących źródeł danych, ustalić, czy zapewniają one wystarczające informacje do przeprowadzenia oceny, i zapewnić, aby dostęp do tych źródeł był możliwy w odpowiednich ramach czasowych.

Istnieje możliwość, że dane ze źródeł administracyjnych, np. z krajowych urzędów statystycznych, zostaną udostępnione podmiotom oceniającym wyłącznie na pewnych warunkach związanych z prywatnością i poufnością danych dotyczących przedsiębiorstw. Warunki dostępu do tych danych muszą być opisane w planie oceny. W razie potrzeby organ zezwalający na dostęp do danych musi zapewnić dostęp do tych danych ekspertom przeprowadzającym ocenę.

Jeżeli wykorzystywane są dane z różnych źródeł, niezwykle istotne jest, aby dane te gromadzono w formacie umożliwiającym konsekwentne dopasowanie zmiennych. Konieczne może być określenie jednakowych identyfikatorów w odniesieniu do jednostek obserwowanych w każdym wykorzystywanym zbiorze danych. Na przykład identyfikatory przedsiębiorstwa i zakładu muszą być jednakowe we wszystkich zbiorach danych, adresy należy gromadzić w formacie umożliwiającym geolokalizację itp. Dokładne źródło identyfikatora może różnić się w zależności od danego państwa członkowskiego. Może on na przykład pochodzić z systemu podatkowego (np. numer VAT) lub być bezpośrednio dostarczony przez instytucje statystyczne (np. SIREN i SIRET we Francji, odpowiednio numer identyfikacyjny przedsiębiorstwa i numer identyfikacyjny zakładu; oba dostarcza Państwowy Instytut Statystyki i Badań Ekonomicznych (INSEE)).

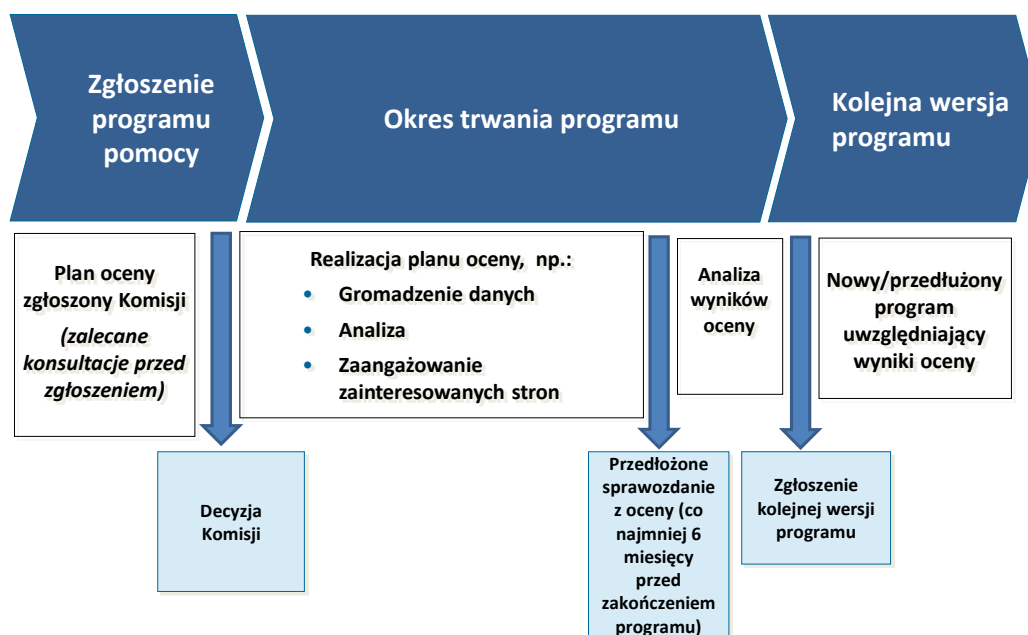
Ocena pomocy państwa może zostać uzupełniona informacjami z badań beneficjentów pomocy lub z wywiadów z podmiotami zarządzającymi programem. Informacje jakościowe tego rodzaju mają charakter subiektywny, a odpowiedzi mogą raczej odzwierciedlać interesy strategiczne beneficjentów, a nie zapewniać prawdziwą ocenę skutków pomocy. Ryzyko to jest szczególnie wysokie, gdy rozmówca zakłada, że pozytywna opinia zwiększy szanse programu na otrzymanie pomocy w przyszłości. Jeżeli jednak podejdzie się do nich z zachowaniem odpowiedniej ostrożności, informacje z badań jakościowych, takich jak wywiady i studia przypadków, mogą stanowić użyteczne źródło dodatkowe oraz mogą pomóc w interpretowaniu wyników oceny.

W każdym przypadku przetwarzania danych osobowych w kontekście oceny zastosowanie ma prawo UE dotyczące ochrony danych osobowych, w szczególności dyrektywa 95/46/WE w sprawie ochrony osób fizycznych w zakresie przetwarzania danych osobowych i swobodnego przepływu tych danych oraz wdrażające ją przepisy krajowe, a także rozporządzenie (WE) nr 45/2001 o ochronie osób fizycznych w związku z przetwarzaniem danych osobowych przez instytucje i organy wspólnotowe i o swobodnym przepływie takich danych.

3.6 Harmonogram oceny

W planie oceny należy podać informacje dotyczące dokładnego harmonogramu oceny, który zostanie określony zgodnie z zatwierdzonym okresem trwania programu, oraz główne etapy, tj. w odniesieniu do gromadzenia danych, przeprowadzania oceny i przedkładania sprawozdania końcowego. Harmonogram może się różnić w zależności od programu i dlatego należy go omówić i uzgodnić z Komisją w odniesieniu do poszczególnych przypadków. Podmiotom zaangażowanym w zarządzanie programami zaleca się, aby umożliwiły przeprowadzenie nieformalnej dyskusji na temat zawartości planu zanim oficjalnie zgłoszą go Komisji.

Aby umożliwić przeprowadzenia oceny proponowanego przedłużenia programu pomocy, należy przedłożyć Komisji sprawozdanie końcowe z oceny w odpowiednim terminie (np. sześć miesięcy przed planowanym zakończeniem programu). Jeżeli nie przewiduje się przedłużenia, sprawozdanie można przedłożyć po zakończeniu programu.



Rysunek 2 – przegląd procesu oceny w przypadku zgłoszonego programu

3.7 Organ przeprowadzający ocenę: zapewnianie niezależności i wiedzy fachowej

Ocena wpływu programów pomocy państwa powinna być obiektywna, szczegółowa, bezstronna i przejrzysta¹¹. Każdą ocenę powinni przeprowadzać w oparciu o wiarygodne metody eksperci, którzy mają odpowiednie i udokumentowane doświadczenie oraz znają metody realizacji tego zadania.

Oceny powinien przeprowadzać organ, który jest niezależny od organu przyznającego pomoc co najmniej pod względem funkcjonowania i który posiada niezbędne i potwierdzone umiejętności oraz personel odpowiednio wykwalifikowany pod kątem prowadzenia takich ocen. Niezależność funkcjonalna podmiotu oceniającego od organu przyznającego pomoc ma zasadnicze znaczenie dla zapewnienia jakości i wiarygodności oceny. Niekoniecznie oznacza to, że należy utworzyć nowy organ albo że ocenę należy zlecić prywatnym podmiotom oceniającym. W zależności od konkretnych organizacji istniejących w każdym państwie członkowskim można byłoby na przykład wykorzystać niezależność i umiejętności organizacji takich jak urzędy statystyczne, banki centralne, trybunały obrachunkowe,

¹¹ Zob. na przykład normy Komisji Europejskiej w zakresie oceny, normy i standardy OECD w zakresie oceny, normy Narodów Zjednoczonych w zakresie oceny oraz niezależna ocena Banku Światowego: zasady, wytyczne i dobra praktyka.

publiczne lub prywatne uniwersytety lub ośrodki badawcze. Decyzję w tym zakresie można podjąć oddzielnie w odniesieniu do każdego programu.

Wczesne zaangażowanie organu przeprowadzającego ocenę, na przykład na etapie opracowywania programu, ma znaczenie dla skutecznego przeprowadzenia oceny. Takie wczesne zaangażowanie zapewnia możliwość oceny programu pomocy państwa w proponowany sposób i gwarantuje gromadzenie niezbędnych danych. W związku z tym w stosownych przypadkach plan oceny powinien być sporządzony przez wyznaczony podmiot oceniający lub co najmniej w ścisłej współpracy z tym podmiotem. Plan oceny powinien również obejmować informacje – nawet jeżeli mają one wyłącznie charakter orientacyjny – dotyczące koniecznych zasobów ludzkich i finansowych, które będą udostępnione na potrzeby przeprowadzania oceny. Szczególne znaczenie mają informacje dotyczące tożsamości i roli każdego głównego eksperta zaangażowanego w ocenę i oszacowanie poziomu zaangażowania tego eksperta.

Plan oceny powinien zawierać dokładny opis organu przeprowadzającego ocenę lub – jeżeli jeszcze go nie wybrano – szczegółowe kryteria, które będą stosowane przy wyborze tego organu, w szczególności kryteria dotyczące niezależności, doświadczenia i umiejętności. W stosownych przypadkach plan oceny powinien uwzględniać istniejące warianty alternatywne. Należy wyraźnie wyjaśnić, dlaczego nie wybrano jeszcze podmiotu oceniającego lub dlaczego podmiot oceniający został wybrany, ale nie uczestniczył aktywnie w sporządzaniu planu oceny. Nawet w takim przypadku plan oceny musi być wystarczająco szczegółowy, aby możliwe było właściwe zbadanie wiarygodności oceny.

3.8 Jawność: ułatwianie zaangażowania zainteresowanych stron

Ocena powinna zostać upubliczniona. Oznacza to, że po zatwierdzeniu zarówno plan oceny, jak i sprawozdanie końcowe z oceny powinny zostać odpowiednio udostępnione poprzez umieszczenie ich w miejscach opisanych w planie oceny, na przykład na stronie internetowej. Komisja również może upublicznić te dokumenty¹².

Jeżeli dane wykorzystywane do oceny są danymi osobowymi lub poufnymi, należy zapewnić poufność w całym procesie oceny – zgodnie z art. 8, 16 i 17 Karty praw podstawowych Unii Europejskiej. Poufność nie obowiązuje jednak w odniesieniu do wyników oceny. W szczególności nie można zawrzeć żadnej klauzuli poufności w umowie dotyczącej oceny, poza: 1. zobowiązaniami dotyczącymi nieujawniania mającymi zastosowanie do danych osobowych lub poufnych; oraz 2. zobowiązaniami dotyczącymi zachowania zgodności z ogólnymi przepisami krajowego prawa statystycznego i tajemnicy statystycznej, takimi jak przepisy związane z prezentacją wyników.

12 Z wyjątkiem tajemnic handlowych i innych poufnych informacji w należycie uzasadnionych przypadkach (komunikat Komisji w sprawie tajemnicy służbowej w decyzjach dotyczących pomocy państwa, C(2003) 4582, Dz.U. C 297 z 9.12.2003, s. 6). Każda publikacja danych osobowych musi zostać przeprowadzona zgodnie z prawem UE w zakresie ochrony danych osobowych, w szczególności z dyrektywą 95/46/WE i przepisami krajowymi wdrażającymi tę dyrektywę oraz z rozporządzeniem (WE) nr 45/2001.

Dane zgromadzone w czasie oceny należy udostępnić w celu umożliwienia powielania wyników lub prowadzenia dalszych badań na warunkach, które nie są bardziej restrykcyjne niż warunki nałożone przez organ przeprowadzający ocenę wstępną.

Organ przyznający pomoc może zapewnić odpowiednie zaangażowanie stosownych zainteresowanych stron, z którymi musi skonsultować się co najmniej raz w czasie realizacji planu oceny. Można na przykład zaprosić zainteresowane strony do omówienia ustaleń oceny wstępnej na podstawie sprawozdania okresowego. Rozwiązania takie należy zawrzeć w planie oceny.

4 Kryteria wyboru dotyczące programów pomocy, które mają zostać poddane ocenie

Zasadniczo każdy program pomocy państwa kwalifikuje się do oceny, jednak – jakkolwiek ocenę uznaje się za dobrą praktykę – to nie zawsze jest ona wymagana na podstawie zasad pomocy państwa. Ocena pomocy państwa powinna zachować proporcjonalny charakter, dlatego z reguły należy przeprowadzać ją w odniesieniu do programów, które mają potencjalne duży wpływ na rynek wewnętrzny i mogą wiązać się z ryzykiem wywołania znaczących zakłóceń, jeżeli w odpowiednim czasie nie zostanie przeprowadzony przegląd realizacji tych programów. W związku z tym stosowne wytyczne w zakresie pomocy państwa dotyczą w szczególności programów pomocy, które: 1) są duże, w tym programów w ramach ogólnego rozporządzenia w sprawie wyłączeń grupowych; 2) mają nowatorski charakter; lub 3) wiążą się z możliwością wystąpienia znaczącej (rynkowej, technologicznej lub regulacyjnej) zmiany w niedalekiej przyszłości, co może wymagać przeprowadzenia przeglądu oceny programu. W poszczególnych wytycznych dotyczących pomocy państwa określono również inne rodzaje programów, którym ocena przyniesie korzyści.

4.1 Duże programy pomocy, w tym programy objęte przepisami ogólnego rozporządzenia w sprawie wyłączeń grupowych

Zgodnie z komunikatem w sprawie unowocześnienia unijnej polityki w dziedzinie pomocy państwa Komisja może wymagać poddania ocenie największych programów pomocy, ponieważ: 1) tego rodzaju programy w największym stopniu wpływają na jednolity rynek, jeżeli nie zostały dobrze zaprojektowane; 2) dzięki ich dużym budżetom można osiągnąć największy przyrost wydajności; oraz 3) duże programy, w których uczestniczy wiele różnych rodzajów beneficjentów, mogą stanowić wystarczające źródło danych na potrzeby oceny.

Niektóre programy pomocy nie muszą jednak podlegać ocenie, jeżeli – mimo swojej wielkości – nie wiążą się z żadnym szczególnym aspektem stwarzającym problemy (np. przypadki typowe, przypadki, w których istnieje duża liczba beneficjentów, ale każdy beneficjent otrzymuje małe kwoty pomocy, oraz przypadki, w których nie istnieje ryzyko znaczących zmian lub nie mogą powstać żadne poważne zakłócenia).

Ponadto w nowym ogólnym rozporządzeniu w sprawie wyłączeń grupowych duże programy pomocy zdefiniowano na podstawie ich budżetu (średni roczny budżet przekraczający 150 mln EUR) oraz przewidziano ocenę niektórych rodzajów pomocy¹³.

Aby nie opóźniać wejścia w życie dużych programów, ale również aby zapewnić, że będą one objęte skuteczną oceną, w ogólnym rozporządzeniu w sprawie wyłączeń grupowych przewidziano zwolnienia z wymogu zgłoszenia na okres wynoszący maksymalnie sześć miesięcy, który Komisja może przedłużyć po zatwierdzeniu planu oceny¹⁴. Plan oceny należy zgłosić możliwie najszybciej, a najpóźniej w ciągu 20 dni roboczych od wejścia w życie programu.

W nowym ogólnym rozporządzeniu w sprawie wyłączeń grupowych przewidziano również zmiany lub kolejne wersje dużych programów podlegających ocenie, które należy zgłosić, jeżeli zmiany te nie mają charakteru wyłącznie formalnego i administracyjnego oraz nie są dokonywane w ramach działań współfinansowanych przez UE.

4.2 Nowatorskie programy pomocy

Definicja „nowatorskiego charakteru” może się różnić w zależności od instrumentów pomocy i państwa członkowskiego. Nowatorski charakter będzie zasadniczo rozpatrywany pod kątem charakteru programu pomocy lub rynków, dla których program jest przeznaczony, np. rynków wschodzących, na których zmiany rynkowe są na bardzo wczesnym etapie. Takie programy mają potencjał kształtowania sektorów przemysłu w trwały i zasadniczy sposób. Zakres zarówno korzyści, jak i zakłóceń jest zatem szczególnie duży. Taki nowatorski charakter może obejmować na przykład wprowadzenie nowego mechanizmu zdolności wytwórczej do sektora energetycznego, pomoc na rzecz nowych rodzajów technologii lub nowatorski rodzaj wsparcia na rzecz odnawialnych źródeł energii w kontekście pomocy na ochronę środowiska. Ocena nowatorskich programów jest także pomocna dla podmiotów aktualnie opracowujących nowe programy, ponieważ umożliwia im uwzględnienie najnowszych zmian na rynku.

4.3 Programy pomocy, na które oddziałują znaczące przewidziane zmiany

Możliwość wystąpienia znaczących (rynkowych, technologicznych lub regulacyjnych) zmian w niedalekiej przyszłości zostanie oceniona w odniesieniu do każdego przypadku oddzielnie. Takie znaczące zmiany mogą obejmować na przykład przewidywaną zmianę mającego zastosowanie rozporządzenia lub pomoc dla branż szybko rozwijających się, gdy otoczenie rynkowe i dostępne technologie rozwijają się w szybkim tempie. Jeżeli programy nie są dostosowane do skutków tych znaczących zmian, występuje ryzyko, że finansowanie publiczne nie będzie skutecznie wykorzystane (na przykład finansowanie może zostać

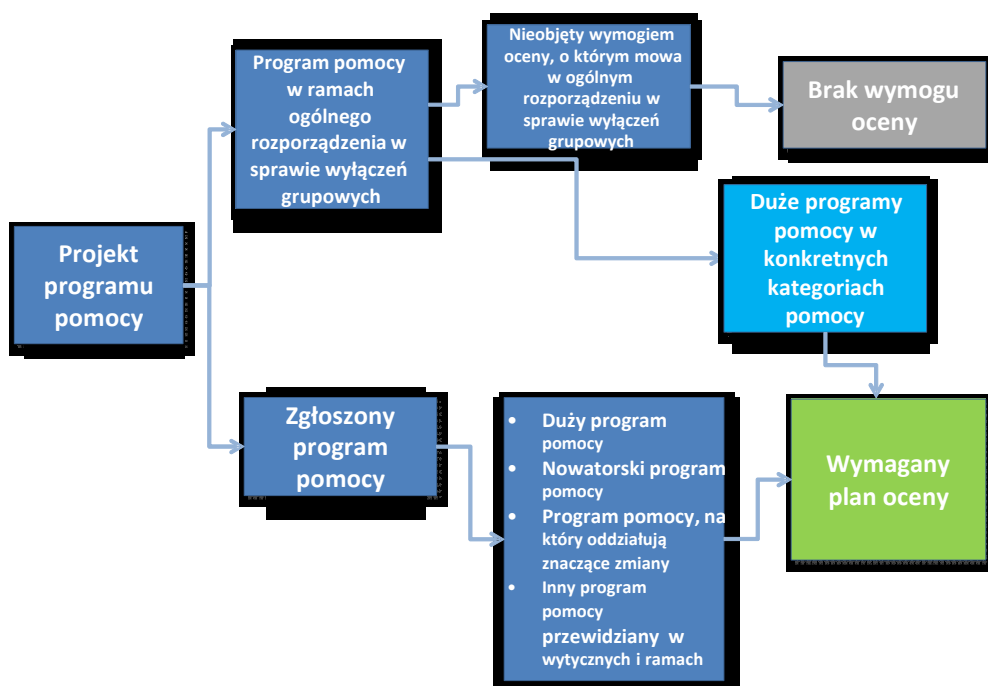
13 Pomocy regionalnej (z wyjątkiem regionalnej pomocy operacyjnej), pomocy dla MŚP, pomocy na dostęp MŚP do finansowania, pomocy na badania, rozwój i innowacje, pomocy na ochronę środowiska (z wyjątkiem pomocy w formie ulg w podatkach na ochronę środowiska na mocy dyrektywy 2003/96/WE) oraz pomocy na infrastrukturę szerokopasmową.

14 W drodze wyjątku Komisja może również postanowić, że ocena nie jest konieczna z uwagi na szczególne cechy danego przypadku.

przyznane na potencjalną „niedoskonałość rynku”, która przestanie istnieć) lub że powstaną znaczące zakłócenia oddziałujące na nowe podmioty na rynku inaczej niż na przedsiębiorstwa działające już na rynku; istnieje również ryzyko, że finansowanie publiczne może doprowadzić do powstania nierównych warunków dla nowych i starych technologii. Przykładami ilustrującymi taką sytuację są zmiana istniejących ram regulacyjnych (na przykład w sektorze łączności elektronicznej), duże wahania cen nakładów lub produkcji (na przykład w przypadku paneli słonecznych) lub wprowadzenie nowej technologii na rynek (na przykład dostępność sieci telefonii ruchomej czwartej generacji w odniesieniu do usług szerokopasmowych). Są to przypadki, w których ocena może być uzasadniona, aby w przyszłych programach można było uwzględnić nowe sytuacje na rynku.

4.4 Inne programy pomocy

W wytycznych dotyczących różnych obszarów pomocy państwa zidentyfikowano również pewne programy pomocy, w przypadku których ocena ma szczególne znaczenie.



Rysunek 3 – wybór programów pomocy do celów oceny

Załącznik I: Dodatek techniczny dotyczący odpowiednich metod określania wpływu przyczynowego

Program pomocy państwa może wywierać wpływ na bardzo różnych poziomach. Zwykle spodziewany jest bezpośredni skutek na poziomie beneficjenta. Aby ocenić stopień wydajności i skuteczności publicznego środka, kluczowe znaczenie ma zrozumienie rozmiarów tego skutku. Ponieważ jednak pomoc ukierunkowuje się na wzajemnie oddziałujące na siebie przedsiębiorstwa na rynkach i w regionach, które konkurują w celu przyciągnięcia działalności gospodarczej, pomoc państwa ma zazwyczaj również skutki pośrednie. Tymi skutkami może być na przykład efekt rozlania w stosunku do innych przedsiębiorstw (np. pozytywny efekt rozlania w ramach badań i rozwoju lub wypieranie inwestycji przez inne konkurujące przedsiębiorstwa) lub efekt przeniesienia (np. przemieszczenie działalności gospodarczej z jednego regionu do innego). Skutki pośrednie powodują zarówno szkody, jak i korzyści wynikające z interwencji państwa w gospodarkę. W związku z tym podczas oceny środka publicznego należy dokonać również oceny rozmiarów skutków pośrednich.

Do pomiaru bezpośrednich i pośrednich skutków polityki zazwyczaj wymagane jest wykorzystanie różnych narzędzi. W ostatnich dziesięcioleciach nastąpił znaczny rozwój metod i technik oceny bezpośredniego wpływu polityk na beneficjentów. Techniki te przedstawiono szczegółowo w dalszej części niniejszej sekcji. Niestety w niewielu okolicznościach da się zastosować te techniki w odniesieniu do oceny wpływu pośredniego, wywieranego przez program pomocy na przedsiębiorstwa lub regiony. Do oceny pośrednich skutków programu pomocy państwa zazwyczaj wymagany jest inny rodzaj dowodów niż dowody wykorzystywane w celu oceny bezpośrednich skutków wpływających na odbiorców, a interpretacja opiera się zazwyczaj na teorii ekonomii i na modelowaniu. Przygotowanie dokładnych wytycznych dla tego rodzaju oceny jest trudniejsze, ponieważ muszą być one dostosowane do możliwych i spodziewanych skutków polityki, zarówno pozytywnych, jak i negatywnych. W związku z tym oceny należy dokonać po dokładnej i szczegółowej analizie najbardziej wiarygodnych możliwych skutków pośrednich polityki. Na podstawie tej analizy podmioty oceniające mogą opracować środki oparte na danych jednostkowych pochodzących od beneficjentów nieobjętych pomocą, w szczególności z tego samego regionu, klastra lub branży, a także z regionów sąsiadujących. Powinny one stanowić podstawę oceny pośrednich skutków programu pomocy państwa. W stosownych przypadkach można je uzupełnić danymi makroekonomicznymi, a przede wszystkim starannie dobranymi studiami przypadku.

Pierwszym niezbędnym i kluczowym krokiem jest ocena skutków bezpośrednich. Szczegółowo przeprowadzona ocena pośrednich skutków pomocy stanowi jednak ważny dowód dla oceny bardziej rozległych skutków programu. Jeżeli brak dodatkowych inwestycji ze strony beneficjentów pomocy stanowi w szerokim rozumieniu wskaźnik niepowodzenia polityki, nawet pozytywny skutek nie wystarczy, aby stwierdzić, że polityka spełniła swoje cele. W szczególności jeżeli okaże się, że bezpośredni wpływ pomocy na beneficjentów jest bardzo mały lub nawet żaden, istnieje duże prawdopodobieństwo uznania programu za

niespełniający zakładanego celu, chyba że uda się w przekonujący sposób uzasadnić, że istnieją istotne i korzystne skutki pośrednie. To samo dotyczy odwrotnej sytuacji: nawet jeżeli z oceny wynika, że bezpośrednie skutki pomocy są pozytywne, pozostaje pytanie, czy mogą istnieć negatywne skutki pośrednie, które by je zrównoważyły lub nawet przeważały.

Ponadto nie zawsze można łatwo i jasno oddzielić skutki bezpośrednie od pośrednich. Może się zdarzyć, że przedsiębiorstwo zainwestowało więcej (domniemany skutek bezpośredni), gdyż jego własne inwestycje wyparły inwestycje przedsiębiorstw konkurencyjnych (pośredni skutek wzajemnego oddziaływania). Przedsiębiorstwo może też inwestować więcej, gdyż spodziewa się efektu rozlania i inwestycji ze strony innych przedsiębiorstw. Ponadto każdy z tych skutków mógł być również wywołany przez pomoc lub zwyczajnie przez samo przyznanie pomocy. Kwestię prawdopodobnego wystąpienia, kierunku i spodziewanych rozmiarów skutków pośrednich należy szczegółowo omówić w ramach oceny skutków bezpośrednich. Teoria ekonomii łącząca skutki pośrednie z pomocą powinna być wyraźnie określona, zaś dodatkowe informacje, które mogą służyć za dowody popierające tę teorię, powinny stanowić integralną część oceny¹⁵.

Wnioskowanie przyczynowe

Wpływ przyczynowy pomocy stanowi różnicę między wynikami osiągniętymi w przypadku przyznania pomocy a wynikami osiągniętymi w przypadku jej braku. Wynik w sytuacji przyznania pomocy występuje w przypadku przedsiębiorstw, które otrzymały pomoc. Wynik w przypadku braku pomocy mierzy się jednak tylko w odniesieniu do przedsiębiorstw, które nie otrzymały pomocy. Z definicji nie bada się, jakie byłyby wyniki, gdyby przedsiębiorstwom, które otrzymały pomoc, jej nie przyznano. W celu oszacowania skutków pomocy w odniesieniu do beneficjentów pomocy konieczne jest więc opracowanie *scenariusza alternatywnego*, tj. ustalenie prawdopodobnego scenariusza oddającego prawdopodobną sytuację odbiorców pomocy w przypadku nieotrzymania pomocy. Należy w tym celu znaleźć *grupę kontrolną*, tj. grupę przedsiębiorstw jak najbardziej podobną pod każdym względem – poza samym faktem otrzymania pomocy – do grupy firm, która otrzymała pomoc.

Jakość grupy kontrolnej ma kluczowe znaczenie dla wiarygodności oceny. Zazwyczaj przedsiębiorstwa otrzymujące pomoc różnią się swoimi cechami charakterystycznymi od przedsiębiorstw, które jej nie otrzymują. Mogą one na przykład prowadzić działalność na uboższym obszarze o mniejszym potencjale rynkowym, mieć mniejszy dostęp do kredytów, być bardziej lub mniej wydajne, mieć projekt do przeprowadzenia lub go nie mieć itp. W związku z tym proste porównanie beneficjentów z podmiotami niebędącymi beneficjentami może odzwierciedlić ten stan rzeczy bardziej niż sam wynik polityki.

15 Chociaż w niniejszym dokumencie skupiono się na bezpośrednich skutkach pomocy, fakt, że pomoc może mieć również skutki pośrednie, nakłada określone wyzwania analityczne w zakresie oceny skutków bezpośrednich; należy w związku z tym przyłożyć szczególną wagę do skutków wzajemnych oddziaływań rynkowych.

Podstawową kwestią związaną z dokonaniem wiarygodnej oceny jest upewnienie się, czy systematyczna różnica między beneficjentami pomocy państwa a podmiotami niebędącymi beneficjentami (tak zwany efekt selekcji) nie obciąża wyników. W ostatnich dziesięcioleciach opracowano szereg metod mających rozwiązać tę kwestię. Wybór metody zależy od ocenianej polityki i od dostępnych danych. Ponadto każda metoda ma swoje ograniczenia i jest wiarygodna jedynie przy pewnej liczbie założeń. Można zwiększyć wiarygodność oceny, dokładnie ustalając i omawiając te ograniczenia. W niniejszym załączniku technicznym przedstawiono najodpowiedniejsze metody, skupiając się na aspektach praktycznych i podkreślając, jak ważna jest dobrze określona strategia¹⁶.

A. Eksperymenty losowe

Określenie odpowiedniej grupy kontrolnej ma kluczowe znaczenie dla uzyskania dobrych (tj. nietendancyjnych) szacunków dotyczących skutku polityki. Najkorzystniejsza jest sytuacja, w której nie występuje efekt selekcji, ponieważ beneficjentów wybiera się losowo¹⁷. Nie występuje wówczas systematyczna różnica między beneficjentami a podmiotami niebędącymi beneficjentami z wyjątkiem samej pomocy, więc różnice w wynikach można przypisać polityce.

Losowa selekcja beneficjentów pomocy spotyka się jednak czasami z krytyką, gdyż jest sprzeczna z celem wielu programów, według których należy wybierać możliwie najlepszych beneficjentów pomocy na podstawie obiektywnych kryteriów. W pewnych okolicznościach istnieje jednak możliwość włączenia elementów losowości do kwalifikowalności lub zachęt beneficjentów do uczestnictwa. Na przykład można ustalić stały budżet dla danego programu. Jeżeli zapotrzebowanie na wsparcie ze strony podmiotów wnioskujących przekracza budżet i charakteryzują się one prawie takimi samymi cechami, można dokonać próby zastosowania losowości. Inny przykład to losowe udostępnianie ewentualnym odbiorcom pomocy różnych poziomów informacji dotyczących programu.

Projekty pilotażowe zapewniają dodatkowe możliwości losowej alokacji pomocy. W przypadku innowacyjnej polityki można zalecić wpierw ocenę programu pilotażowego prowadzonego na mniejszą skalę. Program może być mniejszy, a losowy wybór beneficjentów – łatwiejszy. Inna możliwość zakładałaby stopniową rozbudowę programu, na przykład określając 25 % losowo wybranych przedsiębiorstw jako kwalifikowalne w pierwszym roku, a następnie zwiększając ich liczbę do 50 %, 75 % i 100 % odpowiednio w drugim, trzecim i czwartym roku programu (lub, alternatywnie, rozpowszechniając program wśród coraz większego grona odbiorców). W przypadku nowej polityki okres rozbudowy jest w wielu wypadkach konieczny ze względów administracyjnych.

16 Niniejszy załącznik zawiera szybką i nietechniczną prezentację ekonometrycznych metod oceny polityki. W prezentacji też zapożyczono wiele elementów z Givord (2010). Inne bardzo dobre prezentacje można znaleźć w Imbens i Wooldridge (2009) oraz Angrist i Pischke (2008).

17 Na przykład eksperymenty losowe od dziesięcioleci stanowią jedyną dopuszczalną metodę oceny skutków przyjmowania leków i poddawania leczeniu.

Pomysły takie mogą lepiej pasować do wdrażania całkowicie nowych programów lub wprowadzania znacznych zmian w istniejących programach. Zapewne dość trudno jest wprowadzić randomizację kwalifikowalności (bezpośrednio lub pośrednio) w odniesieniu do kontynuacji istniejącego programu. Nie oznacza to jednak, że nie można posłużyć się eksperymentami losowymi w ramach ich oceny. W szczególności można wybrać losowo beneficjentów do wariantów programu, które są potencjalnie bardziej wydajne, ukierunkowane lub wprowadzają mniejsze zakłócenia. Na przykład w przypadku programu obejmującego dotacje można losowo proponować nowo wprowadzony program pożyczek.

B. Metody quasi-eksperymentalne

Chociaż eksperymenty losowe stanowią najlepszy możliwy sposób oceny skutków polityki, nie zawsze możliwe jest ich wdrożenie. W celu oceny skutków polityki z perspektywy *ex-post* opracowano inne metody. Ich wspólnym celem jest wykorzystanie zewnętrznych różnic środowiska, w którym przedsiębiorstwa prowadzą swoją działalność, aby doprowadzić do sytuacji bardzo zbliżonych do eksperymentów (tzw. eksperymenty naturalne lub quasi-eksperymenty).

Ogólnie identyfikacja eksperymentu naturalnego lub quasi-eksperymentu stanowi wyzwanie w przypadku oceny *ex post*. Szczegółowa analiza struktury polityki może jednak pociągać za sobą analizę występowania wystarczającej zmienności zewnętrznej. W stosownych przypadkach można skorygować początkową konfigurację i wprowadzić więcej elementów, co pozwoli na określenie skutków polityki.

Kontrola pod kątem obserwowalnych różnic

Jak wyjaśniono powyżej zazwyczaj istnieją znaczące różnice między beneficjentami pomocy i podmiotami niebędącymi beneficjentami. Przy porównywaniu wyników tych dwóch grup, tj. beneficjentów pomocy i podmiotów niebędących beneficjentami, konieczne jest uwzględnienie tych różnic.

Wiele różnic w cechach charakterystycznych daje się zazwyczaj zaobserwować. Najpowszechniejszym sposobem uwzględnienia różnic jest wykorzystanie regresji liniowej. W regresji liniowej dąży się do kontroli wpływu obserwowalnych cech charakterystycznych na wyniki. W tej metodzie zakłada się, że występuje zależność liniowa między wynikiem, na przykład inwestycjami w badania i rozwój, a innymi cechami charakterystycznymi przedsiębiorstwa, na przykład sektorem, wiekiem, wielkością itp., w tym również przyznaniem pomocy. Można postrzegać regresję liniową jako aproksymację liniową bardziej złożonych zależności¹⁸. Regresje liniowe można postrzegać jako techniki ogólnego przeznaczenia i stosować w wielu różnych kontekstach oceny.

Jako alternatywę dla regresji liniowej można zastosować techniki dopasowania. Celem technik dopasowania jest łączenie w pary każdego z beneficjentów z innym

¹⁸ Ponadto można zastosować wzajemne oddziaływanie cech charakterystycznych (na przykład sprzedaży i sektora) i wprowadzać funkcje tych cech (na przykład kwadraty zmiennych).

przedsiębiorstwem, które „wygląda” bardzo podobnie, ale nie otrzymało pomocy. Jeżeli chodzi o stosowane w dopasowaniu elementy obserwowalne, mogą to być cechy charakterystyczne przedsiębiorstwa lub szacowane prawdopodobieństwo otrzymania pomocy (technika „propensity score matching”). Dopasowanie może być przydatne do kontroli elementów obserwowalnych w kontekście ważnej strategii empirycznej.

Obydwie techniki – prosta regresja liniowa i dopasowanie – mają jednak swoje wewnętrzne ograniczenia. Obydwie są ważne jedynie przy założeniu tzw. warunkowej niezależności. Przy tym założeniu wymagane jest, aby po uwzględnieniu wpływu obserwowalnych cech charakterystycznych wynik był od nich niezależny. W praktyce oznacza to, że zazwyczaj każda zmienna, mająca wpływ zarówno na wyniki, jak i na selekcję, jest obserwowalna (i jest uwzględniana w formie właściwej funkcji). Jeżeli jest inaczej, sam fakt uczestnictwa przedsiębiorstwa odzwierciedla jego niektóre (nieobserwowane) cechy, które motywują jego działalność. Zarówno przy pomocy regresji liniowej, jak i dopasowania nie uda się uzyskać wiarygodnej oceny. Jeżeli na przykład przedsiębiorstwo posiada „obiecujący projekt”, wpływa to zarówno na prawdopodobieństwo, że będzie wnioskowało o pomoc (i że ją otrzyma), jak i na prawdopodobieństwo, że przedsiębiorstwo to odniesie sukces rozwijając swoją działalność. Brak uwzględnienia tej kwestii wpłynie wyniki.

W szczególności w przypadku dopasowania porównywanie wyników beneficjenta z wynikami dopasowanego „bliźniaka”, który nie otrzymał pomocy, pozwala uniknąć efektu selekcji jedynie w przypadku, gdy przyznanie pomocy nie jest związane z nieobserwowanymi zmiennymi, które również wpływają na wynik. W rzeczywistości takie założenie rzadko się sprawdza. Zmierzenie wszystkich zmiennych, które mają wpływ na fakt wnioskowania o pomoc lub jej uzyskania, rzadko jest możliwe. Ponadto przy wdrażaniu technik dopasowania przedsiębiorstwa otrzymujące pomoc oraz przedsiębiorstwa, które pomocy nie otrzymują, muszą mieć bardzo podobne obserwowalne cechy charakterystyczne. Jeżeli dopasowane przedsiębiorstwa są rzeczywiście pod każdym obserwowalnym względem podobne do siebie, przyczyna, dla której niektóre przedsiębiorstwa otrzymują pomoc, a inne nie, pozostaje z definicji niezaobserwowana. W związku z tym samo istnienie całkowicie kompletnego zbioru danych zawierającego wiele obserwowalnych cech charakterystycznych nie może stanowić uzasadnienia dla ważności oceny opartej na dopasowaniu lub prostszej zwykłej regresji liniowej.

Natomiast ewentualne uzasadnienie wykorzystania do oceny techniki dopasowania lub prostszej regresji liniowej opiera się na fakcie, że niezaobserwowane przyczyny, które wyjaśniają kwalifikowalność lub przyznanie pomocy, nie mają bezpośredniego ani pośredniego wpływu na wyniki (po skontrolowaniu pod kątem cech obserwowalnych). Aby ocena opierająca się na prostej technice dopasowania lub regresji liniowej była wiarygodna, należałoby mieć pewność, że zbiór przedsiębiorstw, które nie otrzymały pomocy, został określony zewnętrznym. Wymaga to, by po skontrolowaniu obserwowalnych cech charakterystycznych nie pozostał żaden niezaobserwowany czynnik wyjaśniający kwalifikowalność lub przyznanie pomocy, który mógłby w sposób bezpośredni lub pośredni wpłynąć na wyniki. Na ogół w przypadku dopasowywania przedsiębiorstw, które są w

jednakowym stopniu kwalifikują się do pomocy, to ostatnie kryterium nie zostanie spełnione. Jeżeli na przykład wszystkie przedsiębiorstwa są kwalifikowalne, w przypadku przedsiębiorstw, które otrzymują pomoc na inwestycje, istnieje znacznie wyższe prawdopodobieństwo, że posiadają projekt, niż w przypadku przedsiębiorstw, które pomocy nie otrzymały (w innym wypadku również wnioskowałyby o pomoc i ją otrzymały). Ogólnie przedsiębiorstwa posiadające projekt mają większe szanse na rozwój pod względem sprzedaży lub zatrudnienia, lecz nie wiąże się to z pomocą i wobec tego dopasowanie na podstawie obserwowalnych cech charakterystycznych nie pozwoli rozdzielić tych dwóch kwestii (chyba że mierzy się istnienie porównywalnego projektu inwestycyjnego).

W wielu sytuacjach stosowanie założenia warunkowej niezależności skazane jest na niepowodzenie. W celu uwzględnienia niezaobserwowanej selekcji może być zatem konieczne wdrożenie w badaniu innych technik niż tylko regresja liniowa czy dopasowanie.

W dalszej części niniejszej sekcji bardziej szczegółowo przedstawiono najpowszechniej stosowane metody oceny wpływu polityki w tym kontekście, tj. różnice w różnicach, nieciągły model regresji (RDD), zmienne instrumentalne (IV). Wiarygodność tych metod opiera się na różnych założeniach, przy czym najlepszy wybór zazwyczaj wynika z kontekstu polityki i dostępności danych. W niniejszej prezentacji omówiono zalety i słabe strony każdej z technik. Poza istotnym wyjątkiem przedstawionych powyżej randomizowanych badań kontrolnych (RTC), nie istnieje żadna technika, która przewyższałaby pozostałe techniki pod każdym względem. Dokonując wyboru techniki należy kierować się dokładną analizą kontekstu danego środka i dostępnymi danymi.

Warto w tym miejscu podkreślić, że możliwość ustalenia skutków polityki nie zależy od stosowania konkretnej techniki ekonometrycznej, tylko od tego, czy grupa kontrolna jest wybierana zewnątrz, a więc od jakości scenariusza alternatywnego. W związku z tym jakość badania oceniającego będzie w dużym stopniu zależała od tego, na ile przekonująco badacz jest w stanie określić zewnętrzną grupę kontrolną. W przypadkach, w których mogą się utrzymywać resztkowe błędy systematyczne, istotne jest szczegółowe omówienie tych błędów, w tym ich źródła, kierunku oraz prawdopodobnego rozmiaru ich wpływu na wyniki.

a. Różnica w różnicach

Przesłanki i identyfikacja

Jak wyjaśniono wcześniej, jest mało prawdopodobne, aby zwykłe porównywanie beneficjentów, a nawet dobrze dobranej grupy podmiotów niebędących beneficjentami prowadziło do wiarygodnej oceny. Wynika to stąd, że nie można wykluczyć występowania niezaobserwowanych różnic między tymi dwoma grupami, co prowadzi do utrzymującej się różnicy w wynikach nawet w sytuacji braku pomocy. Ponadto zwykłe porównywanie wyników beneficjentów przed przyznaniem pomocy i po jej przyznaniu również może doprowadzić do fałszywej oceny. Nie pozwala to na odróżnienie skutków pomocy od skutków wywołanych innymi czynnikami, które również wpłynęły na wynik dwóch grup,

takich jak ogólne tendencje gospodarcze, zmiany w otoczeniu regulacyjnym lub rosnące koszty pracy.

Połączenie obu podejść może jednak umożliwić ocenę związku przyczynowego pomocy: jest to technika różnicy w różnicach (ang. *difference-in-difference*). W sumie polega ona na przeanalizowaniu różnic w wynikach między przedsiębiorstwami w czasie. Wcześniejsze różnice mogły wynikać z innych czynników niż pomoc państwa. Jedynie zmiana w tych różnicach („różnica w różnicach”) może zostać przypisana pomocy. Innymi słowy w metodzie tej porównuje się różnice między wynikami działalności beneficjentów i grupy kontrolnej zarówno *przed* udzieleniem pomocą, jak i *po* jej udzieleniu, a następnie przypisuje się pomocy zmianę w różnicy. Metoda jest skuteczna, jeżeli z czasem zarówno na beneficjentów, jak i na grupę kontrolną wpływają inne czynniki, które wpływają w taki sam sposób również na wyniki działalności. Można więc stwierdzić, że pomoc jest jedynym odpowiednim czynnikiem, który stanowi wyjaśnienie obserwowanej zmiany w wynikach działalności beneficjentów w stosunku do grupy kontrolnej.

Decydujące jest tu założenie, że różnice między beneficjentami i grupą kontrolną są stałe w czasie i że obie grupy identycznie reagują na takie same zmiany (odstępstwa od średniej) w trakcie danego okresu. W praktyce założenie to może zawieść. Jeżeli na przykład beneficjenci są przedsiębiorstwami bardziej podatnymi na zagrożenia, to istnieje większe prawdopodobieństwo, że pogorszenie koniunktury gospodarczej i ogólny klimat biznesowy będą miały na nie większy wpływ. W związku z tym w skład grupy kontrolnej muszą wchodzić również przedsiębiorstwa podatne na zagrożenia. W sumie wybór grupy kontrolnej jest kluczem do wiarygodności metody. Identyfikacja nie polega na wykorzystaniu techniki różnicy w różnicach, która tylko środkiem realizacji technicznej, ale na odpowiednim doborze grupy kontrolnej.

Jeżeli podmioty niebędące beneficjentami same podjęły decyzję, że nie będą występować z wnioskiem o pomoc, należy zachować szczególną ostrożność w tworzeniu grupy kontrolnej. Można oczekiwać, że ubieganie się lub nieubieganie o pomoc jest związane z dochodami, jakie wiążą się z otrzymywaniem pomocy. W związku z tym istnieją podstawy, aby sądzić, że oczekiwane wyniki przedsiębiorstw, które nie ubiegają się o pomoc (pod względem zatrudnienia, wydajności, obrotu itp.), różnią się od oczekiwanych wyników beneficjentów. Jeżeli na przykład wszystkie przedsiębiorstwa ubiegające się o pomoc otrzymują pewną pomoc, to jedynymi kwalifikującymi się przedsiębiorstwami, które nie ubiegają się o pomoc, są przedsiębiorstwa bez projektu (przy założeniu, że koszt ubiegania się jest niski). Wynik działalności takich przedsiębiorstw prawdopodobnie będzie nie tylko gorszy w ujęciu bezwzględnym, ale też porównywalnie gorszy z upływem czasu. Natomiast lepsze przedsiębiorstwa realizują projekty i rozwijają się. Nie można oczekiwać, że zatrudnienie, wydajność lub obrót będą kształtować się na równoległych poziomach, i podwójne różnicowanie na ogół nie rozwiązuje problemu.

W związku z tym przedsiębiorstwa w grupie kontrolnej, które nie skorzystały z pomocy, muszą zostać wybrane z powodów, które nie mają żadnego wpływu na pomiary wyników.

Nie mogły one wybrać się same i dobrowolnie podjęły decyzję o nieuczestniczeniu. Najbardziej przekonująca konfiguracja ma miejsce w przypadku, gdy nieuczestniczenie powiązane jest z brakiem kwalifikowalności, co stanowi efekt naturalnego eksperymentu. W takim przypadku brak kwalifikowalności prawdopodobnie nie jest wynikiem niezaobserwowanych czynników, które również miały wpływ na wyniki. W skład grup kontrolnych mogłyby wchodzić na przykład przedsiębiorstwa usytuowane w regionach, który nie kwalifikują się już do pomocy (jeżeli kwalifikowalność ta nie jest powiązana z ich własnym wynikiem działalności, a raczej z zewnętrznym zdarzeniem).

Wdrażanie

Z technicznego punktu widzenia technika różnica w różnicach może być wdrażana w ramach modelu regresji liniowej albo w z zastosowaniem dopasowywania. W pierwszym przypadku grupa kontrolna zostaje wybrana niezależnie od obserwowalnych cech charakterystycznych i w związku z tym jest ogólnie porównywalna z całą grupą beneficjentów pomocy. Następnie uwzględnia się obserwowane różnice w modelu klasycznej regresji liniowej. W drugim przypadku grupa kontrolna utworzona zostaje z przedsiębiorstw, które są porównywalne pojedynczo z każdym przedsiębiorstwem otrzymującym pomoc w próbie opartej o zaobserwowane czynniki. Wynik każdego przedsiębiorstwa jest porównywany z wynikiem najbardziej porównywalnych przedsiębiorstw i wyniki są agregowane. Te dwie metody stanowią dwa różne sposoby rozpatrywania obserwowanych różnic, ale nie ma żadnej zasadniczej różnicy pod względem identyfikacji związku przyczynowego polityki.

W zależności od okoliczności interesujące może być porównanie zmienności wyników beneficjentów i grupy kontrolnej przed pomocą. Jeżeli wyniki regularnie zaczęły się różnić przed faktycznym udzieleniem pomocy, to istnieje prawdopodobieństwo, że grupa kontrolna i grupa beneficjentów różnią się z przyczyn niezwiązanych z pomocą i stosując tę metodę, nie uzyska się wiarygodnej oceny związku przyczynowego polityki. Nie jest to rygorystyczny test wiarygodności założenia: taki test nie istnieje. Jest to jednak co najmniej przydatna pierwsza kontrola spójności.

Jeżeli istnieje szereg potencjalnych grup kontrolnych, które są *a priori* wiarygodne, można wykorzystać dodatkowe metody i badania odporności. Pierwsza i najbardziej naturalna kontrola odporności polega na wdrożeniu szeregu estymatorów różnicy w różnicach i porównaniu wyników. Ponadto można również wykorzystać te różne grupy kontrolne do skonstruowania bardziej wiarygodnego oszacowania. Wyobraźmy sobie program ukierunkowany na MŚP w konkretnym regionie. Dwie potencjalne grupy kontrolne obejmują przedsiębiorstwa niezaliczane do MŚP w tym regionie lub MŚP w regionie sąsiadującym. Żadne w tych przedsiębiorstwach nie zdecydowało dobrowolnie, aby nie ubiegać się o pomoc – przedsiębiorstwa te po prostu nie kwalifikowały się. Jednak żadna z tych grup kontrolnych nie jest idealna: na większe przedsiębiorstwa w tym samym regionie prawdopodobnie inaczej wpływają ogólne tendencje gospodarcze, natomiast MŚP w regionie sąsiadującym mogą podlegać innym regionalnym wstrząsom. Zamiast wybierać pomiędzy tymi dwoma ewentualnymi estymatorami techniki różnicy w różnicach, można je połączyć i wdrożyć potrójny estymator różnicy (DDD): zaczynając od „klasycznej” techniki różnicy w różnicach

między MŚP i podmiotami niebędącymi MŚP w danym regionie, można uzyskać taka samą różnicę w różnicach na podstawie regionu sąsiadującego, aby zlikwidować różnice w wynikach między MŚP i podmiotami niebędącymi MŚP w regionie z pomocą¹⁹. Zamiennie można systematycznie dokonywać prób zbudowania syntetycznej grupy kontrolnej, składającej się z MŚP z kilku regionów sąsiadujących i podmiotów niebędących MŚP z tego samego regionu, w celu lepszego odtworzenia wzoru wyniku w odniesieniu do beneficjentów przed uzyskaniem pomocy (w celu uzyskania szczegółowych informacji zob. Abadie, Diamond i Hainmuller, 2010 r.).

Wnioskowanie statystyczne

Poza starannym opracowaniem i wyborem grupy kontrolnej należy w szczególności zająć się kwestią wnioskowania statystycznego. Pojęcie wnioskowania statystycznego w tym kontekście odnosi się do pytania, czy skutki, które oszacowano, są faktycznie znaczące. Znaczenie statystyczne to inna kwestia niż znaczenie ekonomiczne. To drugie odnosi się do wielkości oszacowanych skutków w porównaniu z innymi odpowiednimi parametrami z punktu widzenia teorii ekonomicznej. Jak wyjaśniono wcześniej, znaczenie ekonomiczne ma kluczowe znaczenie. Dyskusja zasadniczo ma jednak sens jedynie wtedy, gdy ze statystycznego punktu widzenia skutki są oszacowane wystarczająco precyzyjnie, tj. można wykluczyć, że nie ma żadnego skutku.

Istnieją podstawy, aby sądzić, że bezpośrednie wnioskowanie statystyczne na podstawie standardowych założeń (takich jak założenie homoskedastyczności i założenie braku autokorelacji) prawdopodobnie będzie powodowało zawyżanie znaczenia statystycznego skutków²⁰.

Pierwszy problem dotyczy grupowania danych. Jeżeli zarówno grupa kontrolna, jak i grupa beneficjentów są bardzo jednolite (nawet gdy różnią się między sobą), to na wszystkie przedsiębiorstwa w każdej grupie prawdopodobnie wpłyną podobne odstępstwa od średniej (wstrząsy). W statystyce oznacza to, że wartość błędu ma wspólną część składową. W przypadku gdy wariancja tej wspólnej części składowej jest duża w porównaniu z różnicą w zaobserwowanych wynikach poszczególnych przedsiębiorstw, wynik będzie obciążony błędem systematycznym. W przypadku dwóch okresów i dwóch grup problem ten może być szczególnie poważny i na granicy identyfikacji: nie można oddzielić skutku wstrząsów występujących w obrębie każdej grupy od wpływu polityki. Problem nie musi być tak poważny, jeżeli grupy nie są tak jednolite. Zawsze jednak niezbędne jest zastanowienie się na obecnością wspólnych wstrząsów w jednolitych podzbiorach wchodzących w skład grup.

19 Przykładem może być regionalny program dla MŚP mający na celu tworzenie nowych miejsc pracy. Wyobraźmy sobie, że pod koniec programu okazuje się, że pod względem tworzenia miejsc pracy wyniki działalności MŚP w tym regionie były o 20 % lepsze niż wyniki dużych przedsiębiorstw w tym regionie. Jeżeli w porównywalnym regionie sąsiadującym (gdzie nie udzielono żadnej pomocy) MŚP również miały lepszy wynik działalności niż duże przedsiębiorstwa (przypuśćmy, że lepszy o 15 %), to wpływ pomocy można oszacować na około 5 %.

20 Kwestię tę w znacznym stopniu podkreślono w kontekście techniki różnicy w różnicach, ale takie same problem mogą się pojawić w przypadku innych technik, o których mowa w niniejszym dokumencie.

Jeżeli na przykład popyt ma lokalny charakter, zwykle niezbędne będzie dokonanie korekty pod kątem klastrowej struktury wartości błędu na poziomie lokalnym. To samo miałyby zastosowanie do branż lub sektorów.

Drugi problem pojawia się, gdy stosuje się dane panelowe. Wartości błędu większości danych dotyczących przedsiębiorstw, takich jak zatrudnienie, wydajność i inwestycje, są zazwyczaj automatycznie skorelowane. Oznacza to, że odstępstwa od średniej w jednym okresie prawdopodobnie się utrzymają w następnym okresie. Zlekceważenie tej kwestii prowadzi do zawyżenia precyzji oszacowania skutków i do odrzucenia – znacznie częściej niż należałoby to robić – hipotezy, że polityka nie miała żadnego wpływu. Bertrand, Duflo, and Mullainathan (2004 r.) wykazali, że problem ten może mieć poważny charakter.

b. Zmienne instrumentalne

Przesłanki i identyfikacja

Metoda zmiennych instrumentalnych („IV”) jest klasyczną metodą rozwiązywania problemu endogeniczności zmiennych objaśniających. Ponieważ korzystanie z pomocy może być uważane za endogeniczną zmienną objaśniającą wyniki działalności przedsiębiorstw w kontekście regresji liniowej, naturalne jest użycie zmiennych instrumentalnych w celu oceny skutku pomocy.

Zmienna jest endogeniczna, gdy jest skorelowana z nieobserwowanym elementem, który także wyznacza wynik. Wyobraźmy sobie na przykład próbę identyfikacji wpływu przyznanej pomocy państwa na zatrudnienie w przedsiębiorstwach w drodze dokonania regresji zatrudnienia na udział w programie i inne obserwowane elementy. Wyobraźmy sobie, że program pomocy ukierunkowano na przedsiębiorstwa osiągające słabe wyniki, które prawdopodobnie będą miały do czynienia z trudnymi warunkami rynku lokalnego. Warunki rynkowe nie mogą być obserwowane przez podmiot oceniający i dlatego nie mogą być bezpośrednio kontrolowane. Jeżeli jednak zmienna nie jest uwzględniona, efekt dotacji zostanie prawdopodobnie niedoszacowany przez podmiot oceniający w związku z endogenicznością udziału w programie. To, czy przedsiębiorstwo ma do czynienia z korzystnymi, czy z trudnymi warunkami rynkowymi, ma to wpływ zarówno na udział w programie, jak i na zatrudnienie, tj. udział w programie jest skorelowany z wartością błędu wyjaśniającą zatrudnienie. Wpływ warunków rynkowych na udział w programie oznacza, że nie można przypisać całkowitej korelacji między działem w programie a zatrudnieniem wpływowi przyczynowemu pomocy.

Istnieją jednak także inne czynniki, które wyjaśniają udział w programie, ale nie wyjaśniają zatrudnienia. Na przykład, jak w pracy Criscuolo et al (2012 r.), położenie geograficzne może decydować o kwocie środków dostępnych na potrzeby programu w regionie. Ponadto wykaz regionów objętych programem z czasem podlega zmianom. Jeżeli budżet programu w odniesieniu do danego regionu zmieni się w czasie z powodów zewnętrznych (np. na skutek spadku średniego PKB na mieszkańca w UE), będzie to miało wpływ na udział w programie, ale nie na warunki lokalnego rynku przedsiębiorstwa. Zmiana w zatrudnieniu, która związana

jest ze zmianą zewnętrzną w zakresie programu, nie jest powiązana z warunkami rynku lokalnego. Koncentrując się na tej „części” zmiennej udziału w programie można oddzielić faktyczny wpływ udziału na zatrudnienie w przedsiębiorstwie od zakłóceń warunków lokalnego rynku. Taka jest logika zmiennych instrumentalnych.

W odniesieniu do oceny pomocy państwa zmienna instrumentalna jest zmienną, dzięki której można wyjaśnić fakt otrzymywania pomocy, ale która nie ma bezpośredniego wpływu na inne nieobserwowane wyznaczniki wyniku, którego pomiaru należy dokonać. Zmienne instrumentalne pozwalają zatem skoncentrować się na udziale w programie bez zakłócenia wynikającego ze skutków wyboru. Do celów ilustracyjnych można postrzegać logikę zmiennej instrumentalnej w następujący sposób²¹. Na pierwszym etapie dokonuje się regresji udziału w programie w odniesieniu do wszystkich zmiennych zewnętrznych, w tym zmiennych instrumentalnych. Na drugim etapie zmienną uczestnictwa (zmienna wskazująca, czy otrzymano pomoc) zostaje zastąpiona uczestnictwem, jak przewidziano na etapie pierwszym: takie przewidywane uczestnictwo nie jest skorelowane z nieobserwowanym elementem, który również określa wynik.

Kwestie związane ze słabymi instrumentami

Zmienna instrumentalna jest zmienną, dzięki której można wyjaśnić fakt otrzymywania pomocy, ale która nie ma bezpośredniego wpływu na inne nieobserwowane wyznaczniki wyniku, którego pomiaru należy dokonać. Taka prosta i klasyczna definicja kryje jednak szereg praktycznych trudności. Istnieją badania mające na celu kontrolę spójności instrumentów w przypadku gdy stosuje się więcej instrumentów niż jest to ściśle niezbędne do identyfikacji modelu. Nie istnieją jednak żadne badania wiarygodności instrumentów. Główny cel badania, w którym wykorzystuje się zmienne instrumentalne, polega na wyjaśnieniu, dlaczego można założyć, że każdy pojedynczy instrument jest nieskorelowany z nieobserwowanymi wyznacznikami wyników działalności przedsiębiorstw, czy to zatrudnieniem, czy wydajnością, obrotem, inwestycjami itd. Takie wyjaśnienia, oparte zarówno na argumentach ekonomicznych, jak i elementach rzeczywistych, są niezbędne do zbadania wiarygodności oceny. Nie są one jednak wystarczające, zwłaszcza gdy stosuje się szereg instrumentów.

Omówienie jakości zmiennych instrumentalnych powinno obejmować słabe instrumenty, tj. instrumenty słabo skorelowane ze zmienną wyniku. Jeżeli zmienne instrumentalne są słabo skorelowane ze zmienną endogeniczną, szacunki prawdopodobnie będą nieprecyzyjne. Można w takiej sytuacji pokusić się o dodanie większej liczby zmiennych instrumentalnych. Powszechnie wiadomo, że wprowadzając dużą liczbę zmiennych, można odzyskać w wystarczającym stopniu wstępną zmienną w celu uzyskania znaczących statystycznie wyników. Jednocześnie dwuetapowa metoda najmniejszych kwadratów naturalnie przybliża

21 W praktyce dwuetapowa metoda najmniejszych kwadratów wdrażana jest w jednym kroku z bardzo dobrze znanych powodów związanych z wnioskowaniem statystycznym.

do obciążonego błędem systematycznym zwykłego oszacowania najmniejszych kwadratów²². Należy wyraźnie odnieść się do ewentualności takiego błędu systematycznego w każdej ocenie przy wykorzystaniu metody IV. W szczególności należy rozwiązać problem wiarygodności – nie tylko ich wiarygodności indywidualnej, ale także ich wspólnej egzogeniczności.

Szczególny przypadek pojawia się wówczas, gdy zakłada się, że zmienna endogeniczna jest automatycznie korelowana. Przy założeniu, że źródło endogeniczności występuje tylko równocześnie, można wtedy wykorzystać poprzednie wartości zmiennych instrumentalnych. Należałoby jednak wówczas zastanowić się nad dokładną wiarygodnością tego podejścia. Jeżeli na przykład występuje automatyczna korelacja zmiennych objaśniających, może to mieć także miejsce w przypadku zmierzonego wyniku. Wtedy zmienne opóźnione są również endogeniczne. Ogólniej rzecz ujmując, jeżeli automatyczna korelacja zmiennych objaśniających jest bardzo duża, założenie egzogeniczności może zawieść. Jeżeli jest niewielka, można uciec się do zastosowania wielu opóźnień (i ewentualnie przyszłych wartości), ryzykując wystąpienie problemu związanego z zastosowaniem wielu słabych instrumentów opisanych powyżej. W sumie działanie na przeszłych wartościach może być wiarygodną strategią, ale należy je stosować z zachowaniem ostrożności.

Ogólnie rzecz biorąc, w celu uniknięcia opisanych powyżej problemów, usilnie zaleca się stosowanie jedynie małej liczby przekonywujących instrumentów. Wtedy jednak niezbędne jest również wykazanie, że instrumenty pozwalają dobrze przewidywać endogeniczną zmienną objaśniającą²³.

Różnice w oszacowaniach dwuetapowych: model wyboru Heckmana (1979 r.)

Jeżeli zmienna endogeniczna jest zmienną badania (symulowany udział), pierwsza regresja dwuetapowej metody mniejszych kwadratów może być postrzegana jako liniowy model prawdopodobieństwa dla badanego prawdopodobieństwa. Liniowy model prawdopodobieństwa jest aproksymacją liniową. Jednak w określonych przypadkach badane prawdopodobieństwo, nawet ograniczając się do kwalifikujących się przedsiębiorstw, może być niskie. W związku z tym aproksymacje liniowe mogą być zbyt mało dokładne, aby umożliwiły skuteczną aproksymację badanego prawdopodobieństwa i pozwoliły skoncentrować się na uzyskanych wynikach, które to stanowią przedmiot zainteresowania.

22 Bardzo interesującą dyskusję dotyczącą błędów systematycznych utworzonych przez słabe instrumenty można znaleźć w omówieniu obciążeń statystycznych Bounda, Jeagera i Baker (1995 r.) w opracowaniu Angrista i Kruegera (1991 r.). Ponadto oszacowania zmiennej instrumentalnej są obciążone w skończonym zakresie. W związku z tym nawet w przypadku dużych zbiorów danych w celu zapewnienia widocznego znaczenia statystycznego błędy systematyczne nieasymptotyczne nadal mogą być istotne.

23 Może to mieć formę obliczeń statystyk Fishera pierwszego kroku regresji. Im wyższe obliczenie Fishera, tym mniejsze prawdopodobieństwo, że instrumenty są słabe. Stock, Wright i Yogo (2002 r.) zaproponowali badanie formalne. W odniesieniu do jednego instrumentu jest to na przykład niezbędne, aby statystyka Fishera pierwszego kroku regresji była wyższa niż 10.

Istnieje kilka sposobów rozwiązania tej kwestii. Wszystkie polegają na zastąpieniu liniowego modelu prawdopodobieństwa dla badanego prawdopodobieństwa funkcją nieliniową²⁴.

Klasyczne podejście polega na badaniu problemu oceny w kontekście modelu wybierania (Heckman, 1979 r.). W ramach tego podejścia traktuje się wpływ wyboru na obserwowane elementy i udział jako problem pominiętej zmiennej w regresji liniowej wyniku²⁵. Istnieje szereg wariantów tej metody, na przykład oszacowanie całego modelu na podstawie maksymalnego prawdopodobieństwa lub instrumentalna analiza udzielania pomocy na podstawie przewidywalnych wartości równania wyboru²⁶.

Bardzo istotne jest jednak zastanowienie się nad identyfikacją, a w szczególności nad doborem zmiennych. Nie wystarczy zastosowanie takich samych zmiennych na obu etapach oszacowania, nawet jeśli wyniki są dostatecznie precyzyjne.²⁷

Wiarygodne jest tylko oszacowanie modelu wyboru przy pomocy tak zwanej zmiennej wykluczenia. Wykluczona zmienna to zmienna, która wyjaśnia wybór beneficjenta pomocy, ale nie wynik. Nie wystarczy jednak usunięcie zmiennej z głównego równania, aby ją dodać do wykazu zmiennych objaśniających w równaniu wyboru. Wręcz przeciwnie – ta zmienna wykluczenia musi wyjaśniać wybór, ale nie ma wpływu na wynik, który ktoś próbuje wyjaśnić. Jest w swojej istocie bardzo zbliżona do wiarygodnej zmiennej instrumentalnej. Wybór takiej zmiennej nie może wynikać z wygody; musi on wynikać z teorii ekonomicznej, struktury instytucjonalnej lub doświadczenia.

c. Nieciągły model regresji

24 Niniejszy rozdział zawiera bardzo krótki opis modeli wybierania w tym rozdziale. Bardziej kompletną prezentację można znaleźć w opracowaniu Heckmana (1979 r.) i głównie Wooldridge'a (2002 r.), rozdział 17.

25 Ta pominięta zmienna jest różnicą w warunkowym oczekiwaniu wyniku w odniesieniu do wybranej próby (tutaj beneficjenci pomocy). Na podstawie niektórych założeń do procesu wyboru beneficjentów pomocy (na przykład model probitowy lub logitowy) różnica ta może być formalnie wyprowadzona (odwrócony wskaźnik Millsa) i jest funkcją parametrów wyboru. W związku z tym wpływ polityki można zidentyfikować przez dodanie pominiętej zmiennej do regresji. Parametry wyboru nie są znane, ale spójne parametry mogą zostać odzyskane w oszacowaniu pierwszego kroku w procesie wyboru. Prowadzi to do procedury oszacowania nazywanej niekiedy „Heckit”. W pierwszej kolejności wymaga to odzyskania parametrów zainteresowania wyborem ze strony beneficjentów pomocy, na przykład specyfikacji probit lub logit. Następnie można odtworzyć spójny estymator wpływu polityki przez dodanie oszacowanego odwróconego wskaźnika Millsa do regresji liniowej. Statystyczne pakiety oprogramowania zwykle są wyposażone w funkcję umożliwiającą przeprowadzenie oszacowania Heckmana.

26 Prezentację wszystkich tych metod można znaleźć w pracy Wooldridge'a (2002 r.), rozdział 17.

27 Jeżeli równanie wyboru nie jest liniowe, odwrócony wskaźnik Millsa nie jest współliniowy z innymi zmiennymi objaśniającymi, nawet jeśli pierwsze równanie obejmuje jedynie podzbiór tych zmiennych wyjaśniających. Ponadto w teorii model ten już został zidentyfikowany. W związku z tym odwrócony wskaźnik Millsa bardzo często nie pokazuje wystarczającej różnicy, co prowadzi do bardzo nieprecyzyjnych oszacowań. Jednak szczególnie w przypadku dużych prób oszacowanie może prowadzić do znaczących wyników. Niemniej jednak gdy wszystkie zmienne w modelu wyboru znajdują się także w głównym równaniu, model ten jest jedynie identyfikowany na skutek braku liniowości szczególnej formy parametrowej.

Nieciągły model regresji (RDD) jest najnowszym dodatkiem do zestawu narzędzi do przeprowadzania działań związanych z oceną²⁸. W ciągu ostatniej dekady odniósł on duży sukces w społeczności akademickiej, głównie dzięki swojej prostocie. W metodzie tej wykorzystuje się istnienie zmiennej, która ma nieciągły wpływ na prawdopodobieństwo, na które wpływa polityka. W kontekście programów pomocy państwa przydatnych może być kilka rodzajów nieciągłości. Pierwszym z nich są granice geograficzne: kwalifikowalność programów może być powiązana z dokładnymi granicami administracyjnymi, takimi jak miejscowości, regiony NUTS itp. Drugi rodzaj nieciągłości wynika z warunków nałożonych na przedsiębiorstwa korzystające z programów, w szczególności pod względem wieku i rozmiaru.

Weźmy pod uwagę następujący przykład. Wyobraźmy sobie, że projekty prezentowane przez przedsiębiorstwa są oceniane w 100-punktowej skali i przedsiębiorstwa, które uzyskały co najmniej 70 punktów, otrzymują pomoc, natomiast inne przedsiębiorstwa nie otrzymują żadnej pomocy. Przedsiębiorstwo, które uzyskało 71 punktów, ma minimalnie lepszy projekt niż przedsiębiorstwo, które otrzymało 69 punktów. Konsekwencje tej niewielkiej różnicy są jednak zasadnicze: jedno przedsiębiorstwo otrzymuje pomoc, natomiast drugie nie otrzymuje jej wcale. Rezultaty porównania wyników tych dwóch przedsiębiorstw świadczą dobitnie o związku przyczynowym pomocy.

Formalnie nieciągły model regresji wymaga, aby prawdopodobieństwo otrzymania pomocy było nieciągłe, a wszystkie inne zmienne były ciągłe²⁹. Wdrożenie z technicznego punktu widzenia może być zbliżone do wdrażania zmiennych instrumentalnych, w którym wykorzystuje się jako instrument symulację przekraczającą próg. Istnieją jednak dwie główne różnice. Pierwszą z nich jest fakt, że nieciągły model regresji oparty jest na słabszych założeniach. Przede wszystkim nie jest wymagana *a priori* niezależność instrumentu. Na przykład w przypadku przyznawania punktów przedsiębiorstwa z lepszym projektem mogą ubiegać się o większą pomoc niż przedsiębiorstwa z gorszymi projektami. Jedyne wymóg dotyczy tego, aby w pobliżu progu akceptacji prawdopodobieństwo ubiegania się o pomoc nie było nieciągłe. Drugą różnicę stanowi fakt, że szacunki opierają się jedynie na danych dotyczących przedsiębiorstw, znajdujących się bardzo blisko obu stron progu. Zatem słabsze założenia mają swoją cenę: szacunki RDD mają charakter bardziej lokalny niż zwyczajne szacunki dokonane z wykorzystaniem zmiennych instrumentalnych. Jeżeli skutki pomocy różnią się w odniesieniu do przedsiębiorstw znajdujących się dalej od progu, szacunki RDD nie są poprawnym szacunkiem wpływu pomocy na wszystkich jej beneficjentów.

28 Formalny i pełny opis nieciągłego modelu regresji znajduje się w pracy Imbensa et Lemieux (2008 r.).

29 Formalnie można rozróżnić dwa różne nieciągłe modele regresji: model ostry i rozmyty. W modelu ostrym, który w dorumianym sposób jest tu opisany, wszystkie przedsiębiorstwa – i tylko one – są badane powyżej pewnego progu. W przypadku modelu rozmytego nieciągłość jest mniej radykalna: występuje nieciągłość badanego prawdopodobieństwa, ale nie zmienia się to z 0 na 1. W liczbach bezwzględnych, jeżeli chodzi o programy pomocy oparte na warunkach kwalifikujących, model jest ostry jedynie wówczas, gdy traktowanie uważa się za kryterium kwalifikacji. W przeciwnym razie, jeżeli traktowanie polega na otrzymaniu pomocy, mamy do czynienia z modelem rozmytym. Gdy przyznanie pomocy opiera się na przyznawaniu punktów, pod uwagę brane są jedynie przedsiębiorstwa, które składały wniosek, a model jest ostry.

Obszar takich szacunków może budzić obawy, jeżeli odsuwając się od progu akceptacji można spodziewać się dużych rozbieżności skutków. Ponadto polityka może mieć znaczący wpływ na poszczególne przedsiębiorstwa znajdujące się po drugiej stronie granicy. Może mieć to miejsce na przykład w przypadku istotnych skutków przeniesienia. W takim przypadku wykorzystanie nieciągłego modelu regresji na granicy geograficznej nie jest dobrą strategią empiryczną. Ponadto siłą nieciągłego modelu regresji jest skoncentrowanie się na wąskim paśmie wokół obszaru nieciągłości. Jeżeli szerokość pasma jest duża, nie można zakładać, że wpływ innych charakterystyk jest stały. Problem ten zwykle nie jest rozwiązywany w drodze kontroli obserwowanych elementów, która przyjmuje szczególną funkcjonalną formę.

Graficzna kontrola danych może zapewnić komfort w odniesieniu do wiarygodności założeń stanowiących podstawę nieciągłego modelu regresji. Przede wszystkim należy skontrolować trzy rzeczy. Po pierwsze, należy sprawdzić, czy rzeczywiście na progu przyznawalności pomocy występuje nieciągłość. Po drugie, należy upewnić się, czy mierzone wyniki są nieciągłe w tym samym momencie i czy nigdzie nie występuje inna nieciągłość tego samego rodzaju. Po trzecie, konieczne jest również sprawdzenie, czy w innych parametrach związanych z wynikiem nie występuje żadna nieciągłość, w tym skłonność do ubiegania się o pomoc.

Ponadto nieciągłości mogą być tworzone świadomie w celu umożliwienia oceny programów. W szczególności rozbudowa polityk może być wykorzystana do tworzenia nieciągłości oraz do określenia wpływu prowadzonej polityki.

C. Oszacowanie strukturalne

W niektórych przypadkach można pójść dalej i skonfrontować model teoretyczny, dotyczący np. inwestycji przedsiębiorstw, z danymi w celu uzyskania kluczowych parametrów. Podejście to różni się jakościowo od podejść prezentowanych powyżej. W oszacowaniu strukturalnym wykorzystuje się w pełni określony teoretyczny model zachowania przedsiębiorstw. Oszacowanie umożliwia więc odtworzenie parametrów określających zachowanie przedsiębiorstwa. Pozwala to oszacować jak najdokładniej wyznaczniki indywidualnego zachowania przedsiębiorstw oraz umożliwia przeprowadzenie symulacji skuteczności innych narzędzi. Jednak oszacowanie strukturalne jest na ogół bardziej wymagające pod względem zasobów i danych, a także pod względem założeń.

Nie można podać szczegółowych wytycznych dotyczących oszacowania strukturalnego, ponieważ identyfikacja, oszacowanie i wnioskowanie statystyczne muszą być uzyskane indywidualnie. Niemniej jednak ogólne wytyczne przedstawione powyżej nadal mają zastosowanie. Po pierwsze, model teoretyczny musi pasować do najważniejszych danych rynkowych. Po drugie, kwestie dotyczące nieobserwowanych charakterystyk i wyborów muszą być otwarcie i poprawnie rozwiązane.

D. Dodatkowe uwagi metodologiczne

Niejednorodność skutków traktowania

Poprzednie sekcje dotyczyły głównie oszacowania przeciętnego wpływu traktowania na badany podmiot. Sam nagłówek sugeruje, że skutki pomocy różnią się w zależności od beneficjenta. Niejednorodność ta może mieć wiele przyczyn i wiele konsekwencji. Pierwszą konsekwencją może być fakt, że jeżeli pomoc jest bardzo skuteczna w odniesieniu do jednych przedsiębiorstw, ale dużo mniej skuteczna dla innych, średni wpływ pomocy na podmiot jej podlegający może być nieistotny ze statystycznego punktu widzenia. Brak istotnego ze statystycznego punktu widzenia wpływu pomocy nie oznacza, że pomoc nie wpłynęła na żadne z przedsiębiorstw. Z perspektywy politycznej średnie wyniki dotyczące programu są bardziej interesującym pierwszym wskaźnikiem. Próba zrozumienia wyznaczników tej niejednorodności jest jednak równie ważna w odniesieniu do opracowywania lepszych programów. Pozwala to skupić się wyłącznie na przedsiębiorstwach, dla których pomoc jest najbardziej skuteczna i w najmniejszym stopniu zakłóca konkurencję.

W związku z tym, na ile to możliwe, skutki pomocy powinno się szacować w odniesieniu do różnych rodzajów przedsiębiorstw, takich jak małe przedsiębiorstwa i duże przedsiębiorstwa, krótko i długo działające przedsiębiorstwa, przedsiębiorstwa innowacyjne, przedsiębiorstwa z ograniczonym dostępem do kredytów itp.³⁰.

Zakłócenia w przedsiębiorstwach nieotrzymujących pomocy

Pośrednia lub bezpośrednia ocena wpływu programów na podmioty niebędące uczestnikami dostarcza wielu informacji niezbędnych do dokonania oceny pomocy państwa. Pomoc państwa może zakłócać funkcjonowanie rynku poprzez wpływ na podmioty niebędące beneficjentami, na przykład w wyniku transferów wiedzy od beneficjentów, przez ograniczenie względnej konkurencyjności w stosunku do beneficjentów itp.

Ponadto wpływ pomocy na przedsiębiorstwa lub lokalizacje nieotrzymujące pomocy może mieć wpływ na wiarygodność oceny. Na przykład część wpływu pomocy regionalnej może zmaterializować się w postaci nowych perspektyw na granicy: gdy przedsiębiorstwa historycznie ulokowane po „złej” stronie granicy przenoszą swą lokalizację na drugą stronę. Następnie nieciągły model regresji na granicy obejmował będzie głównie skutki wspomnianego przeniesienia i spowoduje ryzyko przeszacowania realnego zagregowanego wpływu polityki. W takiej sytuacji należy posłużyć się inną strategią empiryczną (pozyteczne może okazać się na przykład sprawdzenie odporności oceny w odniesieniu do większych regionów).

30 Innym podejściem jest systematyczne szacowanie wpływu różnych sposobów traktowania w odniesieniu do przedsiębiorstw będących na różnych etapach warunkowej dystrybucji. Istnieje obszerna literatura, w której szacuje się kwantylowe skutki traktowania, począwszy od pracy Abadiego, Angrista i Imbensa (2002 r.). Jest to bardzo przydatne narzędzie ułatwiające zrozumienie rzeczywistego charakteru niejednorodności traktowania. Jest to jednak mniej przydatne z czysto politycznego punktu widzenia, chyba że możliwe jest bezpośrednie ukierunkowanie różnych przedsiębiorstw w zależności od ich pozycji w warunkowej dystrybucji.

E. Dane

Kluczowe znaczenie ma dostęp do właściwych danych mikroekonomicznych umożliwiających przeprowadzenie oceny wpływu. Dane te muszą zachować spójność w odniesieniu do beneficjentów i podmiotów niebędących beneficjentami. Ponadto dane muszą pochodzić z tego samego źródła, oczywiście z wyjątkiem informacji dotyczących samej pomocy. Dane powinny być dostępne na jak najbardziej dokładnym poziomie, chociaż w niektórych przypadkach, na późniejszych etapach mogą być konieczne pewne formy agregacji.

Niezbędne są dane obejmujące wskaźniki rezultatu zarówno w odniesieniu do grupy badanej, jak i grupy kontrolnej, łącznie z czasem, w którym mierzone są wyniki. Ponadto niezbędna jest możliwie największa ilość danych dotyczących czynników potencjalnie wpływających na wyniki i decyzje podmiotów o udziale w programie pomocy. Dane te wykorzystuje się do kontroli różnic między grupami badanymi a grupami kontrolnymi. Na przykład na poziomie przedsiębiorstwa dane te mogą obejmować lokalizację, rozmiar i informacje demograficzne, a także czynniki produkcji.

Najbardziej naturalnym źródłem danych jest administracja, np. dane z bilansu podatkowego lub badania krajowe. Źródła te dostarczają informacji dotyczących lokalizacji i działalności przedsiębiorstwa, a czasem poszczególnych zakładów. Umożliwiają śledzenie inwestycji i poziomu sprzedaży w rozbiciu na rodzaje działalności, a także obliczanie wskaźników finansowych. Interesujące są również duże badania krajowe lub wspólnotowe, takie jak badania wspólnotowe na temat innowacji. Obejmują one dużą i zarazem reprezentatywną próbkę oraz dostarczają uzupełniających informacji na konkretne tematy. Odpowiednim źródłem informacji są również połączone zestawy danych dotyczące pracodawców i pracowników. Zwykle umożliwiają one powiązanie charakterystyki pracy z każdą lokalizacją zakładu. Może to mieć kluczowe znaczenie, gdy przedmiotem zainteresowania jest wymiar geograficzny pracy.

Pomijając wskaźniki rezultatów i charakterystykę odbiorcy, niezbędne są dane dotyczące pomocy i procesu przyznawania pomocy. Informacje te będą zwykle pochodziły od samego organu przyznającego pomoc. Obejmują one dane dotyczące kwoty pomocy i harmonogram przyznawania pomocy beneficjentom. Jednak ogólne dane dotyczące procesu przyznawania pomocy również są szczególnie pomocne. Dane dotyczące wnioskodawców, którym nie przyznano pomocy, są istotne, zwłaszcza gdy proces przyznawania pomocy odbywa się przy użyciu mechanizmu punktacji³¹.

Dostęp do tak poufnych danych jest zwykle regulowany. W związku z tym kluczowe znaczenie ma zapewnienie terminowego dostępu do tych danych całemu zespołowi naukowemu dokonującemu oceny. Ponadto wspomniane źródła administracyjne zwykle są

31 Posiadanie danych dotyczących odrzuconych wnioskodawców jest szczególnie cenne w odniesieniu do badań prowadzonych w oparciu o podejście nieciągłości regresji.

dostępne z pewnym opóźnieniem. Przy opracowywaniu planu oceny należy zatem wziąć pod uwagę dostępność danych.

F. Przykłady

Przykład 1 (Pomoc regionalna): Crisculo i in. (2012 r.)³² dokonali oceny programu selektywnej pomocy regionalnej (ang. Regional Selective Assistance, RSA) w Zjednoczonym Królestwie w latach 1986–2004. W tym okresie w ramach RSA udzielano dotacji uznaniowej przedsiębiorstwom działającym na obszarach znajdujących się w niekorzystnej sytuacji. Był to główny program wsparcia na rzecz przedsiębiorstw w Zjednoczonym Królestwie. Zakres pomocy przyznawanej w ramach RSA był regulowany wytycznymi w sprawie pomocy regionalnej, a w szczególności mapami regionów kwalifikujących się do pomocy („mapy pomocy regionalnej”). Mapy te z czasem ulegały zmianom. Ogólnie rzecz biorąc, kryterium kwalifikowalności regionu jest względna pozycja regionu pod względem PKB na jednego mieszkańca lub pod względem bezrobocia. W rezultacie status regionu może ulec zmianie z powodu jego rozwoju na przestrzeni pewnego czasu lub z powodu zmiany średniej PKB na mieszkańca w UE (np. gdy nowe państwa członkowskie wstąpiły do UE w 1995 r.) . Ponadto wskaźniki wykorzystane do określenia kwalifikowalności z czasem również ulegają zmianie. W związku z tym część zmian dotyczących kwalifikowalności przedsiębiorstw nie zależy od sytuacji samych przedsiębiorstw, ale raczej od wydarzeń mających miejsce poza granicami Zjednoczonego Królestwa lub od zmian w przepisach administracyjnych. Koncentrując się na tej części zmian dotyczących kwalifikowalności i oceniając sposób, w jaki zmiany te wpłynęły na zmiany w działalności inwestycyjnej, zatrudnieniu i wydajności, Criscuolo i in. (2012 r.) są w stanie w przekonujący sposób określić skutki pomocy.³³

Przykład 2 (Wsparcie przedsiębiorstw): Martini i Bondonio (2012 r.)³⁴ zbadali dwa przypadki dotyczące wsparcia przedsiębiorstw – przypadek dotacji inwestycyjnej dostępnej we Włoszech (ustawa 488) oraz przypadek różnych programów dotyczących MŚP w regionie Piemontu. Szczególnie interesująca jest pierwsza ocean. W jej ramach dokonano porównania przedsiębiorstw, których wniosek o pomoc został zatwierdzony (tj. beneficjentów pomocy), z podobnymi przedsiębiorstwami, których wniosek o pomoc odrzucono, ponieważ budżet przeznaczony na pomoc osiągnął limit. Wykorzystanie odrzuconych wnioskodawców w ocenie jest szczególnie przydatne, aby uniknąć obciążenia selekcji błędem systematycznym, który pojawia się przy próbie porównania wnioskodawców z podmiotami niebędącymi wnioskodawcami. Ta grupa przedsiębiorstw przeszła pierwszy test jakości, co oznacza, że przedstawiły wiarygodny projekt inwestycyjny. W związku z tym podzielają ambicje beneficjentów pomocy dotyczące inwestycji w wiarygodne projekty. Z powodu limitów budżetowych (racjonowanie) nie otrzymali jednak pomocy. Różnica w wyniku między

32 Criscuolo, C., R. Martin, H. Overman i J. Van Reenen, 2012 r. „The causal effects of an industrial policy”, CEPR Discussion Papers 8818, C.E.P.R. Discussion Papers.

33 Technicznie rzecz ujmując, Criscuolo i in. (2012 r.) stosują podejście oparte na zmiennych instrumentalnych, jak przedstawiono w niniejszym dodatku technicznym.

34 Sprawozdanie dla DG Regio. A. Martini i D. Bondonio: „Counterfactual impact evaluation of cohesion policy: impact and cost effectiveness of investment subsidies in Italy” (2012 r.).

wnioskodawcami, którzy (prawie) otrzymali pomoc a wnioskodawcami, których wniosek (prawie) odrzucono, dostarcza wiarygodnych oszacowań dotyczących skutków pomocy.

Przykład 3 (Gwarancje kredytowe): Lelarge, Sraer i Thesmar (2010 r.) ocenili skutki programu gwarancji kredytowych we Francji. W ramach programu „SOFARIS” zapewniano ubezpieczenie kredytodawców od ryzyka niewypłacalności kredytobiorców w formie gwarancji. Kredytobiorcy płacą składki ubezpieczeniowe, ale składki te są subsydiowane. Lelarge, Sraer i Thesmar (2010 r.) jednoznacznie określają charakter skutków wyborów. Po pierwsze, przedsiębiorstwa posiadające bardziej rentowne projekty są bardziej skłonne zgodzić się na uiszczenie opłaty związanej z gwarancją. Po drugie, osoby zarządzające programami są skłonne wybierać projekty pożądane ze społecznego punktu widzenia, które w innym przypadku mogłyby nie otrzymać dostępu do prywatnego finansowania. Ogólnie rzecz biorąc, przedsiębiorstwa same wybierają się do programu, a wybór ma miejsce również na etapie przyznawania pomocy. Prawdopodobnie wpływa to na wyniki naiwnych ocen opartych na przykład na klasycznych regresjach liniowych lub porównaniach z najbardziej podobnymi przedsiębiorstwami.³⁵ Faktyczny i instytucjonalny kontekst programu stanowi jednak źródło identyfikacji skutków polityki. Program ustanowiono w późnych latach 80. XX wieku i początkowo obejmował jedynie przedsiębiorstwa działające w branżach produkcji oraz świadczenia usług dla przedsiębiorstw. W 1995 r. publiczne środki finansowe programu zostały zwiększone i nowe branże (przemysł budowlany, handel hurtowy i detaliczny, transport, branża hotelarsko-gastronomiczna, usługi osobiste) stały się kwalifikowalne. Lelarge, Sraer i Thesmar (2010 r.) porównują nowo kwalifikowalne przedsiębiorstwa z przedsiębiorstwami już wcześniej kwalifikowalnymi w celu dokonania oceny wpływu programu na różne wskaźniki, takie jak zadłużenie, zatrudnienie, wzrost kapitału, koszty finansowe i prawdopodobieństwo upadłości. Przedsiębiorstwa znajdujące się w tych dwóch grupach często się różnią. Na przedsiębiorstwa powinny jednak mieć wpływ podobne wstrząsy makroekonomiczne i w związku z tym różnice między nimi nie powinny ulegać zmianie wraz z upływem czasu, z wyjątkiem oczekiwanych efektów samej polityki³⁶.

35 Sytuacja taka ma miejsce, gdy techniki dopasowania – tutaj dopasowanie do siebie najbliższych sąsiadów – nie stanowią lepszego sposobu na rozwiązanie problemów wyboru niż metoda najmniejszych kwadratów. Zgodnie z wcześniejszymi wyjaśnieniami zawartymi w niniejszym dodatku technicznym techniki dopasowania, ogólnie rzecz biorąc, nie są sposobem na rozwiązanie kwestii związanej ze skutkami wyboru, gdy brakuje naturalnych eksperymentów.

36 W praktyce autorzy wdrożyli model wyboru Heckmana ze zmienną wykluczającą na poziomie przedsiębiorstw i z klasyczną strategią IV na poziomie sektora. Bardziej szczegółowe informacje na tych temat metod można znaleźć we wcześniejszej części niniejszego dodatku technicznego.

Przykład 4 (Kreatywny kredyt): Bakhshi i in.³⁷ wykorzystują randomizowane badanie kontrolne, aby ocenić skutki innowacyjnego programu wsparcia na rzecz przedsiębiorstw. Badanie pilotażowe, które rozpoczęło się w Manchesterze w 2009 r., skonstruowano tak, aby bony lub „kreatywne kredyty” mogły być losowo przyznane MŚP wnioskującym o pomoc na inwestycje w kreatywne projekty, takie jak tworzenie stron internetowych, produkcja filmów wideo i kreatywne kampanie marketingowe, w celu sprawdzenia, czy mają one realny wpływ na innowacje. Dzięki kreatywnym kredytom powstały zupełnie nowe stosunki między MŚP a kreatywnymi przedsiębiorstwami; przydzielenie kreatywnych kredytów zwiększało prawdopodobieństwo, że przedsiębiorstwa rozpoczną innowacyjne projekty we współpracy z kreatywnymi przedsiębiorstwami, z którymi wcześniej nie współpracowały, o co najmniej 84 %. Badanie wykazało, że przedsiębiorstwa, którym przyznano kreatywny kredyt, doświadczyły krótkoterminowego wzrostu innowacji i sprzedaży w okresie sześciu miesięcy od ukończenia kreatywnego projektu. Pozytywne skutki nie były jednak trwałe i po upływie 12 miesięcy różnica między grupami, które otrzymały kredyt, a tymi, które go nie otrzymały, przestała być znacząca ze statystycznego punktu widzenia. Sprawozdanie dowodzi, że wyniki pozostałyby niedostrzegalne przy wykorzystaniu zwykłych metod oceny stosowanych przez rząd; w sprawozdaniu wzywa się do szerszego korzystania z randomizowanego badania kontrolowanego przy ocenie polityki w celu wsparcia rozwoju przedsiębiorstwa.

Przykład 5 (Wsparcie w zakresie badań, rozwoju i innowacji): Einiö (2013 r.) analizował wpływ dotacji Tekes na rzecz badań i rozwoju na inwestycje w badania i rozwój, na zatrudnienie i wydajność w okresie od 2000 r. do 2006 r. Tekes jest krajową agencją innowacji odpowiedzialną za większą część wsparcia na rzecz badań i rozwoju w Finlandii. W badaniu wykorzystano regionalne różnice dotyczące potencjalnej wysokości budżetu Tekes na przyznanie wsparcia na badania i rozwój, wynikającej ze zwiększonego finansowania z EFRR w północnej i wschodniej części Finlandii (obszary celu 1). Obszary początkowo określono w ramach negocjacji w sprawie przystąpienia Finlandii w 1995 r. i dokonano tego na podstawie zasady dotyczącej gęstości zaludnienia, według której liczba osób na kilometr kwadratowy nie może przekraczać 8. W związku ze stosunkowo większym budżetem na wspieranie badań i rozwoju, prawdopodobieństwo otrzymania wsparcia było wyższe w obszarach celu 1 w porównaniu z innymi częściami kraju. Spowodowało to regionalną różnicę w traktowaniu przy znacznie większej frakcji przedsiębiorstw wspieranych w obszarze celu 1. Ponieważ alokacja regionów oparta była na wcześniej określonej zasadzie dotyczącej gęstości zaludnienia wynikającej z danych z 1993 r. dotyczących gęstości zaludnienia (a nie na podstawie przewidywanych przyszłych poziomów inwestycji w badania i rozwój lub wyników gospodarczych), w badaniu kontroluje się gęstość zaludnienia z 1993 r., co stanowi skuteczną odpowiedź na obawy dotyczące selekcji regionalnej. W praktyce skutki traktowania szacuje się z wykorzystaniem podejścia opartego na zmiennych

37 Sprawozdanie dla organizacji Nesta, „Creative credits, a randomized controlled industrial policy experiment”, Bakhshi, H., J. Edwards, S. Roper, J. Scully, D. Shaw, L. Morley i N. Rathbone (2013 r.), dostępne pod adresem: http://www.nesta.org.uk/sites/default/files/creative_credits.pdf.

instrumentalnych, w którym wskaźnik dla regionu celu 1 jest użyty jako instrument wejścia do programu. Podejście to pozwala określić wpływ wsparcia wśród przedsiębiorstw, które przystąpiły do programu wsparcia, jako skutek zwiększonego finansowania w obszarze celu 1. Wiarygodność konfiguracji potwierdzono, udowadniając, że było różnic między przedprogramowymi tendencjami wśród przedsiębiorstw, które przystąpiły do programu, i tendencjami w grupie kontrolnej. Einiö (2013 r.) zauważa pozytywny wpływ na inwestycje w badania i rozwój, zatrudnienie i sprzedaż wśród podmiotów uczestniczących, którym przyznano dotację na badania i rozwój w wyniku dodatkowego połączonego wsparcia w finansowaniu na rzecz badań i rozwoju w ich regionie. Chociaż natychmiastowy wpływ na wydajność nie występuje, badania dostarczają dowodów na długoterminowy wzrost wydajności.

Załącznik II: Wykaz możliwych wskaźników rezultatu

Należy zauważyć, że poniższy orientacyjny wykaz przeznaczony jest wyłącznie do celów ilustracyjnych. Faktyczne wskaźniki rezultatu należy wyznaczać zgodnie z celem programu pomocy i oceny.

Bezpośredni wpływ pomocy na szczeblu beneficjentów

	WYMIAR REZULTATU	WSKAŹNIKI REZULTATU
Pomoc regionalna	Pozytywny wpływ	<p>Inwestycja sektora prywatnego odpowiadająca wsparciu publicznemu</p> <p>Wzrost zatrudnienia we wspieranych przedsiębiorstwach</p>
Pomoc na działalność badawczą, rozwojową i innowacyjną	Dodatkowa działalność w zakresie badań, rozwoju i innowacji	<p>Inwestycja sektora prywatnego odpowiadająca wsparciu publicznemu</p> <p>Dodatkowe wydatki na badania, rozwój i innowacje poniesione przez wspierane przedsiębiorstwa</p> <p>Liczba nowo zatrudnionych naukowców we wspieranych przedsiębiorstwach</p> <p>Liczba nowo zarejestrowanych patentów</p> <p>Liczba przedsiębiorstw wspieranych w celu wprowadzenia nowości na rynki</p>
Pomoc na ochronę środowiska	Pozytywny wpływ środowiskowy	<p>Zmniejszona emisja CO₂ w przedsiębiorstwach będących beneficjentami</p> <p>Zwiększenie zdolności wytwarzania energii odnawialnej</p> <p>Zmniejszenie udziału składowanych lub spalanych odpadów</p> <p>Liczba oczyszczonych terenów zanieczyszczonych</p>
	Wcześniejsze przyjęcie norm środowiskowych	<p>Odsetek przedsiębiorstw spełniających nowe normy środowiskowe co najmniej X miesięcy/lat przed ich wejściem w życie [ponieważ wymagany jest co najmniej 1 rok, a pomoc o wyższej intensywności jest dozwolona, jeżeli okres ten wynosi ponad 3 lata]</p>

Pomoc w dziedzinie energii (na infrastrukturę energetyczną)	Zmniejszone zużycie energii	<p>Liczba gospodarstw domowych z lepszą klasą zużycia energii</p> <p>Spadek rocznego zużycia energii pierwotnej przez budynki publiczne</p> <p>Liczba dodatkowych użytkowników energii podłączonych do inteligentnych sieci</p>
	Wsparcie dla energii odnawialnej	Udział produkcji energii z OZE
Finansowanie ryzyka	Pozytywny wpływ	<p>Zwrot osiągnięty w ramach funduszu</p> <p>Lewarowanie inwestycji prywatnych</p> <p>Liczba przedsiębiorstw otrzymujących kapitał wysokiego ryzyka</p>
	Typowanie przegranych	Słabe średnie wyniki jednostek, w których dokonano inwestycji, spowodowane niedociągnięciami w zarządzaniu na zasadach rynkowych/niewystarczającym uczestnictwem prywatnym
	Brak wystarczającego stopnia zróżnicowania	Zbyt małe/ograniczone regionalnie fundusze z ograniczoną możliwością zwrotu, które pozostają nieatrakcyjne dla inwestorów prywatnych
Pomoc na sieć szerokopasmową	Zwiększony zasięg łączności szerokopasmowej	<p>Objęcie zasięgiem dodatkowych gospodarstw domowych przy połączeniu szerokopasmowym o przepustowości co najmniej 30 Mbps</p> <p>Objęcie zasięgiem dodatkowych gospodarstw domowych lub wykorzystanie przez nie połączenia szerokopasmowego o przepustowości co najmniej 100 Mbps</p>
	Wydajność	<p>Koszty inwestycji/pomoc na inwestycje od przyłączenia gospodarstwa domowego (gospodarstwa domowe w zasięgu)</p> <p>Liczba gospodarstw domowych przyłączających się do korzystania z nowych usług</p>
Ratowanie i restrukturyzacja	Pozytywny wpływ	<p>Utrzymanie zatrudnienia i działalności na poziomie właściwym dla przedsiębiorstwa i na poziomie regionalnym</p> <p>Zmiany udziału w rynku i wydajności przedsiębiorstw otrzymujących pomoc</p>

Lotnictwo	Pozytywny wpływ	<p>Liczba przewoźników lotniczych korzystających z portu lotniczego</p> <p>Inwestycja sektora prywatnego odpowiadająca wsparciu publicznemu</p> <p>Zwiększenie produktywności regionalnej lub wartości dodanej brutto (WDB)</p>
	Negatywne skutki	<p>Powielanie infrastruktury lub tras lotniczych przynoszących straty</p> <p>Pogorszenie warunków ruchu w obrębie istniejącej infrastruktury (np. inne porty lotnicze na obszarze ciążenia lub inne środki transportu)</p>

Pośredni wpływ programu pomocy

	WYMIAR REZULTATU	WSKAŹNIKI REZULTATU
Możliwe pozytywne skutki	Zyski makroekonomiczne	Wzrost zatrudnienia Zwiększenie wydajności lub wartości dodanej brutto (WDB)
	Zróżnicowanie gospodarki regionalnej	Liczba branż podlegających różnym kodom NACE
	Zwiększona współpraca pomiędzy jednostkami prywatnymi a jednostkami publicznymi	Liczba przedsiębiorstw współpracujących z instytucjami badawczymi
	Pozytywne efekty zewnętrzne/rozłania	Liczba beneficjentów pośrednich (np. liczba osób trzecich korzystających z obiektu) Zmiany w zatrudnieniu lub działalności innych przedsiębiorstw i regionów (lotnictwo) Liczba mieszkańców korzystających z udoskonalonych środków transportu na obszarze ciężenia
Możliwy negatywny wpływ na konkurencję i handel	Tendencyjność sektorowa	W programie wielosektorowym przeważającą pomoc przyznano jednej branży
	Tendencyjność wobec przedsiębiorstw przynoszących straty lub przedsiębiorstw o niskiej wydajności (zapobieganie likwidacji)	Stosunek przedsiębiorstw o dużej wydajności do przedsiębiorstw o małej wydajności.
	Tendencyjność wobec jednostek zasiedziały	Stosunek starych przedsiębiorstw do nowych przedsiębiorstw
	Umocnienie władzy rynkowej	Zmiana władzy rynkowej przedsiębiorstw dominujących
	Skutki dotyczące lokalizacji	Przeniesienie z biedniejszego regionu do regionu bardziej rozwiniętego
	Dla bezpieczeństwa dostaw	Uzależnienie od wysokoemisyjnych źródeł energii Ocena, czy obawy dotyczące wyłączenia prądu są rzeczywiste i czy nadal występują Przejęcie krajowych rynków energii elektrycznej

	Dla infrastruktury energetycznej	<p>Przejęcie krajowych rynków energii elektrycznej</p> <p>Umocnienie władzy rynkowej jednostki zasiedziałej</p>
	Ratowanie i restrukturyzacja	<p>Zmiany w zatrudnieniu lub działalności innych przedsiębiorstw i regionów</p> <p>Zmiany udziału w rynku i wydajności przedsiębiorstw otrzymujących pomoc</p>
	Lotnictwo	<p>Powielanie infrastruktury lub tras lotniczych przynoszących straty</p> <p>Pogorszenie warunków ruchu w obrębie istniejącej infrastruktury (np. inne porty lotnicze na obszarze ciężenia lub inne środki transportu)</p>

Załącznik III:Glosariusz

Wartość bazowa	Wartość wskaźnika przez podjęciem danej interwencji.
Grupa kontrolna	Analiza scenariusza alternatywnego wymaga znalezienia najbardziej porównywalnego przedsiębiorstwa (przedsiębiorstw) lub grupy kontrolnej, np. grupy przedsiębiorstw, które powinny być jak najbardziej podobne do grupy przedsiębiorstw, które otrzymały pomoc – z tą różnicą, że nie skorzystały one z pomocy.
Scenariusz alternatywny	W celu oszacowania wpływu pomocy na beneficjentów pomocy konieczne jest sporządzenie „scenariusza alternatywnego”, tj. ustalenie prawdopodobnego scenariusza oddającego sytuację, w jakiej mogliby znaleźć się beneficjenci pomocy, gdyby jej nie otrzymali.
Ocena	Systematyczne gromadzenie i analiza informacji na temat programów i projektów, ich celu oraz realizacji; czerpanie wiedzy dotyczącej ich wpływu jako podstawy do wydawania osądów. Z ocen korzysta się w celu zwiększenia skuteczności i podejmowania decyzji dotyczących bieżącego i przyszłego programowania.
Wpływ	Zmiana, którą można w sposób wiarygodny przypisać interwencji. Oznacza samo, co „efekt” interwencji lub „przyczynienie się do zmiany”.
Wskaźnik	Zmienna stanowiąca źródło informacji ilościowych lub jakościowych dotyczących konkretnego zjawiska. Zazwyczaj obejmuje wartość i jednostkę miary.
Metoda	Metody oznaczają rodziny technik oceny i narzędzi służących rozmaitym celom. Zwykle składają się one z procedur i reguł, które zapewniają usystematyzowanie podejścia do badania i spójności w jego przeprowadzeniu. W metodach można położyć nacisk na gromadzenie lub analizę informacji i danych; mogą być ilościowe lub jakościowe; w ich ramach można podejmować próby opisanego, wyjaśnienia, przewidzenia lub dostarczenia informacji niezbędnych do działania. Wybór metody wynika z charakteru interwencji, zadanych pytań ewaluacyjnych oraz funkcji oceny – przyczynowej, poznawczej lub normatywnej itp.
Rezultat	Szczególny wymiar dobrobytu społeczeństwa, który jest motywacją dla realizacji działań w ramach polityk, tj. który docelowo ma być zmieniony poprzez zaplanowane interwencje i wdrożony poprzez realizację polityki. Przykładami są: mobilność w danym obszarze; umiejętności w danym sektorze działalności.
Wskaźnik rezultatu	Wskaźnik opisujący określony wymiar rezultatu, cechę, którą można zmierzyć. Przykładami są: czas potrzebny do przemieszczenia się z punktu A do punktu B z określoną prędkością w kontekście mobilności; wyniki egzaminów w kontekście zwiększenia poziomu umiejętności; udział przedsiębiorstw, którym odmówiono kredytu bez względu na wysokość oprocentowania kredytu w kontekście racjonowania kredytów przez banki.

Załącznik IV: Dokumenty referencyjne

Abadie, A., J. Angrist i G. W. Imbens (2002), „Instrumental Variables Estimates of the Effect of Subsidised Training on the Quantiles of Trainee Earnings”, *Econometrica*, 70(1), s. 91–117.

Abadie, A., A. Diamond i J. Hainmueller (2007), „Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California’s Tobacco Control Programme”, *Journal of the American Statistical Association*, czerwiec 2010 r., tom 105, nr 490.

Angrist, J. i A. Krueger (1991), „Does Compulsory School Attendance Affect Schooling and Earnings”, *Quarterly Journal of Economics*, 106.

Angrist, J. i J. Pischke (2008), „Mostly Harmless Econometrics: An Empiricist’s Companion”, Princeton University Press.

Angrist, J. D. i J. Pischke (2010), „The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics”. *Journal of Economic Perspectives*, 24(2): s. 3–30.

Bakhshi, H., J. Edwards, S. Roper, J. Scully, D. Shaw, L. Morley i N. Rathbone (2013), „Creative credits, a randomized controlled industrial policy experiment”, *Sprawozdanie dla organizacji Nesta*, dostępne pod adresem: http://www.nesta.org.uk/sites/default/files/creative_credits.pdf.

Bertrand, M., E. Duflo i S. Mullainathan (2004), „How much should we trust differences-in-differences estimates?”, *The Quarterly Journal of Economics*, 119, s. 249–275.

Bound, J., D. Jaeger i R. Baker (1995), „Problems with Instrumental Variable Estimation When the Correlation Between the Instruments and the Endogenous is weak”, *Journal of the American Statistical Association*, 90(430), s. 443–450.

Criscuolo, C., R. Martin, H. Overman i J. Van Reenen (2012), „The causal effects of an industrial policy”, *CEPR Discussion Papers 8818*, C.E.P.R. Discussion Papers.

Duflo, E., R. Glennerster & M. Kremer (2007), „Using Randomisation in Development Economics Research: A Toolkit”, *CEPR Discussion Papers 6059*, C.E.P.R. Discussion Papers.

Duflo, E. i M. Kremer (2005), „Use of Randomisation in the Evaluation of Development Effectiveness”, w *Evaluating Development Effectiveness*, pod redakcją O. Feinstein, G. K. Ingram i G. K. Pitman. New Brunswick, New Jersey and London, U.K.: Transaction Publishers, tom 7, s. 205{232}.

Einiö, Elias (2013), „R&D Subsidies and Company Performance: Evidence from Geographic Variation in Government Funding Based on the ERDF Population-Density Rule”, The Review of Economics and Statistics (w przygotowaniu).

Standardy oceny Komisji Europejskiej. Dostępne pod adresem: http://ec.europa.eu/dgs/secretariat_general/evaluation/docs/standards_c_2002_5267_final_en.pdf

Garicano, L. C. Lelarge i J. Van Reenen, (2012), „Firm Size Distortions and the Productivity Distribution: Evidence from France”, CEP Discussion Papers dp1128, Centre for Economic Performance, LSE.

Givord, P. (2010), „Méthodes économétriques pour l'évaluation de politiques publiques”, WPD3E nr G2010-08.

Givord, P., Rathelot, R. i P. Sillard (2013), „Place-based tax exemptions and displacement effects: An evaluation of the Zones Franches Urbaines programme”, Regional Science and Urban Economics, tom 43, wyd. 1, styczeń 2013 r., s. 151–163.

Heckman, J. J. (1979), „Sample Selection Bias as a Specification Error”, Econometrica 47, s. 153–161.

Imbens, G. i J. Wooldridge (2009), „Recent Developments in the Econometrics of Programme Evaluation”, Journal of Economic Literature, 47:1, s. 5–86.

Imbens, G. W., et T. Lemieux (2008), „Regression discontinuity designs: A guide to practice”, Journal of Econometrics, 142(2), s. 615–635.

C. Lelarge, D. Sraer i D. Thesmar (2010), „Entrepreneurship and Credit Constraints: Evidence from a French Loan Guarantee Programme”, NBER Chapters, w: International Differences in Entrepreneurship, s. 243–273, National Bureau of Economic Research, Inc.

Keane, M. P. (2010), „A Structural Perspective on the Experimentalist School”, Journal of Economic Perspectives, 24(2): s. 47–58.

Martini, A. i D. Bondonio (2012), „Counterfactual impact evaluation of cohesion policy: impact and cost effectiveness of investment subsidies in Italy”, sprawozdanie dla Komisji Europejskiej, DG Regio.

Nederlandse Rijksoverheid (2012), „Durf te meten”, Eindrapport Expertwerkgroep Effectmeting, dostępne pod adresem: <http://www.rijksoverheid.nl/documenten-en-publicaties/rapporten/2012/11/23/durf-te-meten-eindrapport-expertwerkgroep-effectmeting.html>

Nevo, A. i M. D. Whinston, (2010), „Taking the Dogma out of Econometrics: Structural Modeling and Credible Inference”. Journal of Economic Perspectives, 24(2): s. 69–82.

Normy i standardy OECD w zakresie oceny. Dostępne pod adresem: <http://www.oecd.org/dac/evaluation/dcdndep/41612905.pdf>

Sims, C. A. (2010), „But Economics Is Not an Experimental Science”. *Journal of Economic Perspectives*, 24(2): s. 59–68.

Stock, J., J. Wright and M. Yogo (2002): „A Survey of Weak Instruments and Weak Identification in Generalised Method of Moments”, *Journal of Business and Economic Statistics*, 20(4), s. 518–29.

United Nations Evaluation Group (2005), „Standards for Evaluation in the UN System”. Dostępne pod adresem: http://www.uneval.org/papersandpubs/documentdetail.jsp?doc_id=22

Wooldridge, J. (2002), „Econometric Analysis of Cross Section and Panel Data”, Cambridge: MIT Press.

Bank Światowy (2003), „Independent Evaluation: Principles, Guidelines and Good Practice”. Dostępne pod adresem: <http://siteresources.worldbank.org/INTDGF/Resources/Evaluation&LearningNote.pdf>